Stacked Data Matrices Workshop

Software Setup

Giuseppe Carteny

24.05.2021

Contents

1 Overview

This workshop is designed to be run on your laptop, and this document summarizes the steps to set up the needed softwares. After following the instructions you should have the softwares properly set up to fully participate in the workshop.

2 R and RStudio Setup

R and RStudio are separate downloads and installations. R is the underlying statistical computing environment. RStudio is a graphical integrated development environment (IDE) that makes using R much easier and more interactive. You need to install R before you install RStudio. After installing both programs, you will need to install some specific R packages within RStudio.

2.1 Windows

2.1.1 If you already have R and RStudio installed

- Open RStudio, and click on "Help" > "Check for updates". If a new version is available, quit RStudio, and download the latest version for RStudio.
- To check which version of R you are using, start RStudio and the first thing appearing in the console indicates the version of R you are running. Alternatively, you can type `sessionInfo()`, which will display your current R version. Go on the CRAN website and check whether a more recent version is available. If so, please download and install it. If you wish to remove an older version you can check here.

2.1.2 If you don't have R and/or RStudio installed

- Download R from the CRAN website, and install it;
- Go to the RStudio download page.

2.2 macOS

2.2.1 If you already have R and RStudio installed

- Open RStudio, and click on "Help" > "Check for updates". If a new version is available, quit RStudio, and download the latest version for RStudio.
- To check which version of R you are using, start RStudio and the first thing appearing in the console indicates the version of R you are running. Alternatively, you can type `sessionInfo()`, which will display your current R version. Go on the CRAN website and check whether a more recent version is available. If so, please download and install it. If you wish to remove an older version you can check here.

2.2.2 If you don't have R and/or RStudio installed

- Download R from the CRAN website, and then double click on the `.pkg` file for the latest R version;
- It might be a good idea to install XQuartz, that may be needed by some R packages;
- Go to the RStudio download page, select the Mac OS installer, and then install it.

2.3 R packages

In order to build a stacked data matrix in R you will need a set of R packages in order to import, reshape, and mutate/recode the data. The tidyverse is an opinionated collection of R packages designed for data science. All packages that are part of it share an underlying design philosophy, grammar, and data structures. The magrittr package provides a set of operators, such as the 'pipe'-like operator `%>%`, which make your code more readable. Although being part of the tidyverse, you will have to load it separately in order to have at hand the full set of operators that the package provides. The haven package enables R to read and write various data formats used by other statistical softwares. As in the case of magrittr, haven is part of the tidyverse, nonetheless you will have to load it separately. The data.table package consists in another package for data manipulation operations (such as subset, group, update, join,...), sometimes considered as an alternative to the tidyverse[1]. Finally, the labelled package provides functions to manipulate metadata as variable labels. This package is particularly useful when working with Stata or SPSS dataset files.

In order to install and load the packages mentioned above, open RStudio and then run the following code:

```r
# Create a string vector containing the name of the required packages # - - - - - - - - - -
want = c("tidyverse", "magrittr", "haven", "data.table", "labelled")

# Load the packages if installed, otherwise install and load them # - - - - - - - - - - - -
have = want %in% rownames(installed.packages())
if ( any(!have) ) { install.packages( want[!have] ) }
junk <- lapply(want, library, character.only = TRUE)

# Prevent scientific notation # - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
options(scipen = 99)

# Remove all the objects from the environment # - - - - - - - - - - - - - - - - - - - - - -
rm(list = ls())
```

Congratulations! You are now ready for the workshop. See you on Thursday!

---

[1] If you are interested, you can find a rather comprehensive comparison here.

3 Stata Setup

The stacked data matrix can be realised also using Stata, a renowned software for data manipulation, visualization, statistics, and automated reporting. Stata is a licensed software, therefore you should obtain a license and activation key before installing it. Once obtained the license/activation key check whether you should install Stata/MP, Stata/SE, or Stata/BE.

3.1 Windows

3.1.1 If you already have Stata installed

- If you have a direct Internet connection, type `update query` in Stata, and Stata will tell you if there are updates and what to do next.
- If you don't have a direct Internet connection, check whenever possible the Stata support page.

3.1.2 If you don't have Stata installed

Download:

- Go to the Stata download site;
- Input your username, password and serial number in the prompt. You received these in your email from StataCorp;
- On the next screen, enter your information;
- On the next screen, choose your operating system;
- Run the `.exe` file.

Installation:

- Close other programs, and click on 'Next' to begin the setup;
- Read and accept the license agreement. Click 'Next';
- Enter your name and your company or organization, and choose whether you wish to allow anyone who uses your computer to access Stata or only you;
- Select the type and edition of Stata to install next, and click on Next. **Note that** if you try to install a edition of Stata different from the one you purchased, you will receive an error later on, and Stata will not work on your computer. Then you will have to uninstall and reinstall Stata with the correct settings to resolve the issue;
- Select a Destination Folder. Stata will choose one automatically. If it does not exist, Stata will create it. If you wish to have Stata installed somewhere else, click on Browse to navigate to that location;
- Choose the default location for saving Stata files. You can use each user's Documents folder, or use a fixed folder for all users. We strongly recommend that you use each user's Documents folder;
- After this passage you are ready to install the application. Install it and then click on 'Finish'.

Set up the license:

- Launch Stata;
- You must enter something for all the fields in the *Initialize License* dialog before you can continue. The code and authorization are not case sensitive. If you make a mistake when typing the codes, you will be prompted to try again;
- Enter your name (at least five characters) and your organization (at least five characters);
- Enter (or, better, copy-paste) the *Serial Number*, *Code*, and *Authorization* from the License and Activation Key you received from StataCorp. Click on 'Next' to submit your information;
- Click on 'Finish' to complete the installation and register your copy of Stata.

3.2 macOS

3.2.1 If you already have Stata installed

- If you have a direct Internet connection, type `update query` in Stata, and Stata will tell you if there are updates and what to do next.
- If you don't have a direct Internet connection, check whenever possible the Stata support page.

3.2.2 If you don't have Stata installed

Download:

- Go to the Stata download site;
- Input your username, password and serial number in the prompt. You received these in your email from StataCorp;
- On the next screen, enter your information;
- On the next screen, choose your operating system;
- Click on the `.dmg` file to download the disk image to your Downloads folder;
- Once the disk image has finished downloading, open your Downloads folder and double-click on the `.dmg` file;

Installation:

- Double-click on 'Install Stata'.
- Read the requirements to confirm that your hardware and system is compatible with Stata;
- Read the *Software License Agreement*. Click on 'Agree' to accept the agreement and continue with the installation;
- Confirm which version of Stata you purchased to install. Check which version of Stata you want to install. It must match the version that is on your license and activation Key. With your selection made, click on 'Continue';
- The next screen will confirm where the Installer will install Stata, and it will tell you how much disk space the installation will take. Click on 'Install' to proceed with the installation;
- Once the installation is complete click 'Close' to finish the installation and close the Installer.

Set up the license:

- Launch Stata;
- You must enter something for all the fields in the *Initialize License* dialog before you can continue. The code and authorization are not case sensitive. If you make a mistake when typing the codes, you will be prompted to try again;
- Enter your name (at least five characters) and your organization (at least five characters);
- Enter (or, better, copy-paste) the *Serial Number*, *Code*, and *Authorization* from the License and Activation Key you received from StataCorp. Click on 'Next' to submit your information;
- Next click on 'Finish' to complete the installation and register your copy of Stata.

3.3 StackMe package

The best option for creating a stacked data matrix on Stata is the StackMe package, an evolution of the PTVTools package, developed by Lorenzo De Sio (then European University Institute, now Luiss Rome) and Mark N. Franklin (then European University Institute, now Emeritus at Trinity College Connecticut) originally published on SSC in 2011.

Two steps are required in order to install StackMe.

```
* 1. Install the 'github' Stata package:
net install github, from("https://haghish.github.io/github/")


* 2. Once this is done, you can install in Stata the latest version of StackMe:
github install ldesio/stackme
```

Congratulations! You are now ready for the workshop. See you on Thursday!