**Instructions getting started in R**

This document is a complimentary resource to the main paper (Reference) , and the R script. We have prepared this document to lead beginners through to the point of using the script and the example dataset. When using this document, we refer to notes in the R script. A note is indicated by a *hashtag* ‘#’ at the beginning of the statement. The R script is written to be as intuitive as possible, but it is not automatic. R is a command line program, and each script sequence requires you to press ‘CTRL+Return’ (Windows) or ‘Command+Return’ (Mac OSX).

To begin you should download either *R* or *RStudio* for free. The script is written to be compatible with both options:

<https://www.r-project.org/>

<https://www.rstudio.com/products/rstudio/download/>

Once installed, right-click on the Script; ‘Supplementary Information 2.R’ and choose ‘Open with’ -> *‘R’* , or ‘*R Studio’*. The script will appear in one window, and the outputs will appear in another. Arrange your windows so you can see both. If the command you enter has worked correctly, it will appear in the outputs window followed by a blank line starting with , ‘>’. If a command you enter has not worked correctly, an error will appear in the outputs window. The script we present here is complex, and while you can use it following our instructions we recommend becoming familiar with *R* first. There are many books, online resources, tutorials and videos for this. A good introductory text is ‘The R book” (Crawley 2007). *R* is case-sensitive, and a command line will not work if there is a small error in syntax or symbols. *R* itself and *RStudio* can have updates and may stop working if you do not keep everything updated. There is much documentation online about how to update *R* according to your operating system and unique needs.

***# Install packages***

Packages are collections of *R* functions, data, and compiled code in a well-defined format (Crawley 2007). The directory where packages are stored is called the library. *R* comes with a standard set of packages. Others are available for download and installation for free. Once installed, they have to be loaded into the session to be used.

There are 6 packages required for this analysis. Install each one, entering each command under the #install packages note. Watch the output window for the packages to install to your library, and watch for any indications or errors. Sometimes a package may have ‘dependencies’ which are other packages the package you are installing depends on. We have written the script to also install dependencies here.

***# Load packages from your Library***

After you install the packages, you need to load the package from your library. Anytime you close *R*, you must load the packaged from your library you need again. Load each package from your library by entering the command under this note. Watch the output window to see the library load the packages and any indications it may tell you.

Sometimes a package may need to be updated too. If this is the case, the indications in the output will report this to you. There are resources online to guide you through how to do this depending on your operating system.

***# Set working directory***

In R you must set your working directory for every script and every new project. This simply tells R which folder, or file path in your computer to find the data files you will be using. This will be different for every user. If you move the data files, then the path, or working directory, will need to be changed accordingly. It is wise to choose simple folder names, and have a special working directory for every R project you start. The very first line is where you set your own working directory. This will be your first lesson in remembering that R is case sensitive, and will not work if syntax is incorrect or small typos are present.

For Windows users this will look like;

“C:/R”

For Mac OSX Users;

“/Users/nameofuserscomputer/Desktop/R/”

***# Load dataset***

The script is written to load the example dataset. Note the name of the dataset file, ‘S1 Example Data.txt’. When using your own data, simply change the name of this file to the name of your own dataset. Put the dataset in the folder or working directory you have indicated in the previous step.

We have loaded the dataset and given it a name in R, ‘data’. Anytime you see the name ‘data’ in the script, we are using this dataset we just imported. You can give a name to something or create an object in R by using this same method here ‘data<-’.

***# Explore the data***

Here we have put some helpful commands in to explore the data and get to know what you are working with, try each one, line by line. Notes explain each command.

***# Ready to begin!***

A one-line command requires you to move your cursor to that line and enter the command. A multi-line command will be identifiable because each line beneath the first will be indented and come immediately after the previous lines. In a multi-line command you must highlight the entirety of the command and enter it. All commands are closed with either a bracket ‘)’ or curly braces ‘}’.

Remember, some commands will have no visible output, even though the command has executed properly. If you see the command followed by a blank line beginning with ‘>’, the command has workedand R is ready for the next command. If you see the command followed by a blank line beginning with ‘+’ then you have not entered the complete command, and R is waiting for you to enter the rest of the information.

We have given notes at every stage of the script to tell you what you are doing and any indications or actions you must take. As the paper indicates you will produce four new files, and four to five figures. We walk you though each one. Enjoy!

**Literature Cited**

Crawley MJ (2007) *The R Book*. Wiley.