

**Tutorial**

**Data Envelopment Analysis with dear**



**Version 1.0 (English)**  
**(Noviembre 2018)**

**Vicente Coll-Serrano<sup>(1)</sup>**

**Rafael Benítez<sup>(2)</sup>**

**Vicente J. Bolós<sup>(3)</sup>**

(1) Department of Applied Economics. [Vicente.Coll@uv.es](mailto:Vicente.Coll@uv.es)

Quantitative Methods for Measuring Culture (MC2)

(2) Department of Business Mathematics. [Benitez.Suarez@uv.es](mailto:Benitez.Suarez@uv.es)

(3) Department of Business Mathematics. [Vicente.Bolos@uv.es](mailto:Vicente.Bolos@uv.es)

School of Economics ([www.uv.es/economia](http://www.uv.es/economia))

University of Valencia (Spain)

# Index

	Page
1. Introduction.....	1
2. Download and install R and RStudio.....	1
2.1. R installation.....	1
a) Installing R on Windows.....	2
b) Installing R on Mac.....	2
2.2. RStudio installation.....	3
3. Getting started with RStudio.....	4
3.1. How to create a script.....	4
4. How to create and work with projects in RStudio.....	5
4.1. Creating a project.....	5
4.2. Opening a project.....	6
4.3. Additional information.....	6
4.4. How to create a project. An example.....	6
5. How to install and load deaR.....	7
<b>5.1. Installing deaR.....</b>	<b>7</b>
<b>5.2. Loading deaR.....</b>	<b>8</b>
6. Save the script and close the working session.....	9
<b>7. Data Envelopment Analysis with deaR.....</b>	<b>10</b>
7.1. Importing data into RStudio.....	11
7.2. Adapting the data to deaR.....	16
7.2.1. Getting help from deaR.....	16
<b>7.2.2. read_data() function.....</b>	<b>18</b>
<b>7.2.3. read_malmquist() function.....</b>	<b>20</b>
<b>7.2.4. read_data_fuzzy() function.....</b>	<b>24</b>
<b>7.3. Select and run a DEA model.....</b>	<b>27</b>
<b>7.4. Extracting the main results.....</b>	<b>32</b>
<b>7.5. summary() function.....</b>	<b>34</b>
<b>7.6. Graphical results: The plot() function.....</b>	<b>41</b>
<b>Acknowledgements.....</b>	<b>48</b>

## 1. INTRODUCTION.

**deaR** is a new and versatile R package (free software) that allows to run a wide variety of models based on Data Envelopment Analysis.

This tutorial is written so that non-R users can use **deaR**, but it is not a manual of introduction to R<sup>1</sup>.

If you are an R user, you can directly go to section 5 and 7.

We want **deaR** to become the reference software tool for those researchers, practitioners, teachers, students and users of the Data Envelopment Analysis methodology. For this, any comments and suggestions to improve **deaR** will be really appreciated. We also accept suggestions of DEA models to be considered for being programmed in new versions of **deaR**. In this sense, we like to anticipate that a new version of **deaR** will include models or features not currently covered such as: stochastic DEA, network DEA, Malmquist index (other decompositions and bootstrapping), negative values and extension of undesirable inputs/outputs to other models in which they are not available now, etc.

We will release a **deaR shiny version** (an interactive web app) soon. We will inform you when it available.

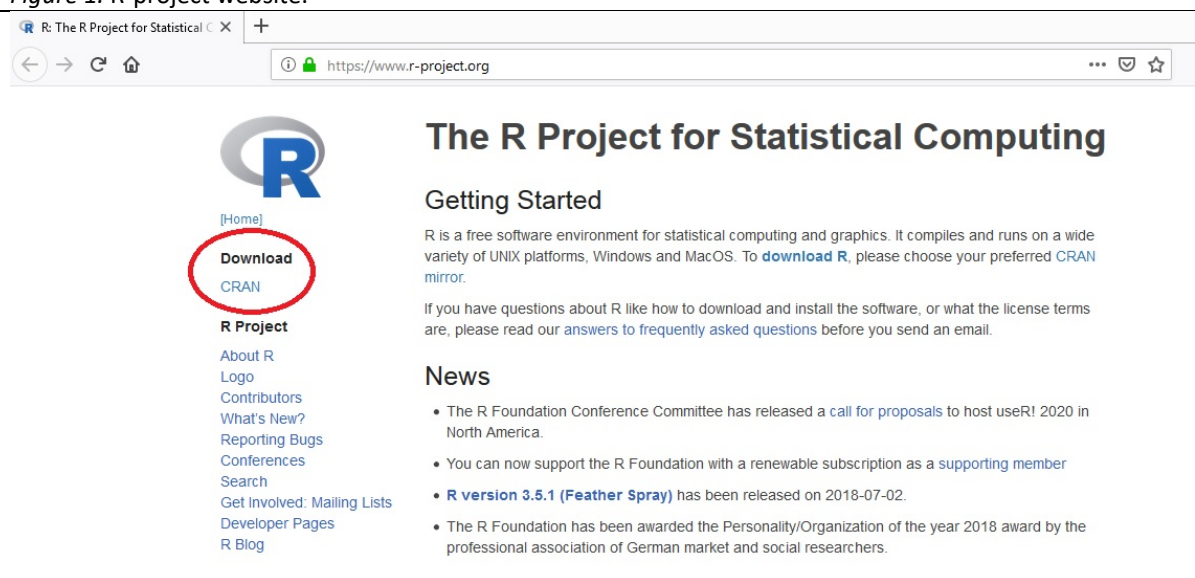
## 2. Download and Install R and RStudio.

To use **deaR**, firstly we have to download and install R and Rstudio.

### 2.1. R installation.

To install R, we go to the R-project website: <http://www.r-project.org> (see Figure 1).

Figure 1. R-project website.



To download R, we click the *CRAN* link and select the “mirror site” nearest to our location.

Now, depending on the operating system of our computer, we select the appropriate option (see Figure 2).

<sup>1</sup>Manual of introduction to R: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>

Figure 2. Versions of R.

---

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

### a) Installing R on Windows.

By clicking the *Download R for Windows* link, we will go to the webpage shown in the Figure below. We click on *install R for the first time* (see Figure 3).

Figure 3. R for Windows.

---

R for Windows

Subdirectories:

<a href="#">base</a> <a href="#">contrib</a> <a href="#">old.contrib</a> <a href="#">Rtools</a>	<p>Binaries for base distribution. This is what you want to <a href="#">install R for the first time</a>.</p> <p>Binaries of contributed CRAN packages (for R <math>\geq</math> 2.13.x; managed by Uwe Ligges). There is also information on <a href="#">third party software</a> available for CRAN Windows services and corresponding environment and make variables.</p> <p>Binaries of contributed CRAN packages for outdated versions of R (for R <math>&lt;</math> 2.13.x; managed by Uwe Ligges).</p> <p>Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.</p>
--	---

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.

Next, we click on **Download R 3.5.1 for Windows** and save the installation file (see Figure 4)

Figure 4. Download R for Windows.

---

R-3.5.1 for Windows (32/64 bit)

[Download R 3.5.1 for Windows](#) (62 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Now, we open the downloaded file (by double-clicking on the file) to install R.

### b) Installing R on Mac.

By clicking the *Download R for (Mac) OS X* link, we will go to the webpage shown in the Figure below. We click on **R-3.5.1.pkg** to download the installation file (see Figure 5).



Figure 5. R for Mac.

---

R for Mac OS X

This directory contains binaries for a base distribution and packages to run on Mac OS X (release 10.6 and above). Mac OS 8.6 to 9.2 (and Mac OS X 10.1) are no longer supported but you can find the last supported release of R for these systems (which is R 1.7.1) [here](#). Releases for old Mac OS X systems (through Mac OS X 10.5) and PowerPC Macs can be found in the [old](#) directory.

Note: CRAN does not have Mac OS X systems and cannot check these binaries for viruses. Although we take precautions when assembling binaries, please use the normal precautions with downloaded executables.


As of 2016/03/01 package binaries for R versions older than 2.12.0 are only available from the [CRAN archive](#) so users of such versions should adjust the CRAN mirror setting accordingly.

R 3.5.1 "Feather Spray" released on 2018/07/05

**Important:** since R 3.4.0 release we are now providing binaries for OS X 10.11 (El Capitan) and higher using non-Apple toolkit to provide support for OpenMP and C++17 standard features. To compile packages you may have to download tools from the [tools](#) directory and read the corresponding note below.

Please check the MD5 checksum of the downloaded image to ensure that it has not been tampered with or corrupted during the mirroring process. For example type `md5 R-3.5.1.pkg` in the *Terminal* application to print the MD5 checksum for the R-3.5.1.pkg image. On Mac OS X 10.7 and later you can also validate the signature using `pkgutil --check-signature R-3.5.1.pkg`

Lastest release:



**R-3.5.1.pkg**  
MD5-hash: 58b3d796d8cf96fb8580c62f46ab64d4  
SHA1: e0c1bfa62a6896d5f4e4511e25d17276d149621  
(ca. 74MB)

R 3.5.1 binary for OS X 10.11 (El Capitan) and higher, signed package. Contains R 3.5.1 framework, R.app GUI 1.70 in 64-bit for Intel Macs, Tcl/Tk 8.6.6 X11 libraries and Texinfo 5.2. The latter two components are optional and can be omitted when choosing "custom install", they are only needed if you want to use the `tcltk` R package or build package documentation from sources.

Note: the use of X11 (including `tcltk`) requires [XQuartz](#) to be installed since it is no longer part of OS X. Always re-install XQuartz when upgrading your macOS to a new major version.

**Important:** this release uses Clang 6.0.0 and GNU Fortran 6.1, neither of which is supplied by Apple. If you wish to compile R packages from sources, you will need to download and install those tools - see the [tools](#) directory.

---

We open *R-3.5.1.pkg* and follow the instructions to install R.

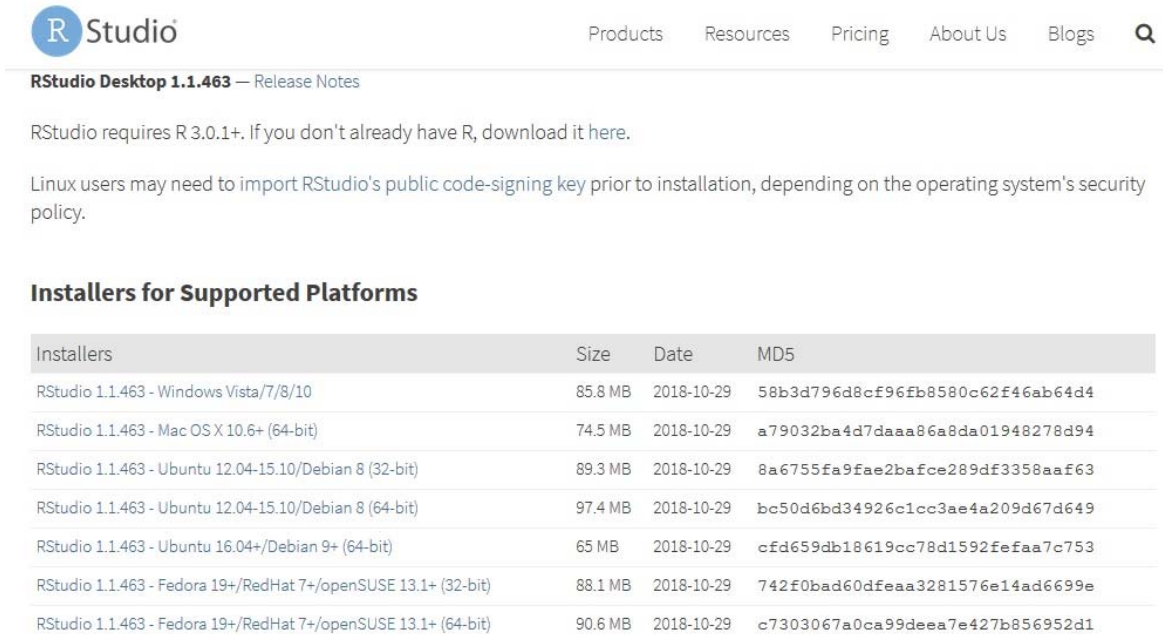
## 2.2. RStudio installation.

Once we have installed R, we download RStudio from the following link (see Figure 6):

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

Figure 6. Download RStudio.

---



R Studio Products Resources Pricing About Us Blogs Q

**RStudio Desktop 1.1.463** — Release Notes

RStudio requires R 3.0.1+. If you don't already have R, download it [here](#).

Linux users may need to import RStudio's public code-signing key prior to installation, depending on the operating system's security policy.

**Installers for Supported Platforms**

Installers	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab64d4
RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-10-29	a79032ba4d7daaa86a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	8a6755fa9fae2bafce289df3358aaf63
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-10-29	bc50d6bd34926c1cc3ae4a209d67d649
RStudio 1.1.463 - Ubuntu 16.04+/Debian 9+ (64-bit)	65 MB	2018-10-29	cf659db18619cc78d1592fefaa7c753
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-10-29	742f0bad60dfeaa3281576e14ad6699e
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-10-29	c7303067a0ca99deea7e427b856952d1

---


To download the executable file, we select the option according to our operating system:

- RStudio 1.1.463 - Windows Vista/7/8/10
- RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)

First, we save the file. Then, we open it to install RStudio. We follow the installation instructions.

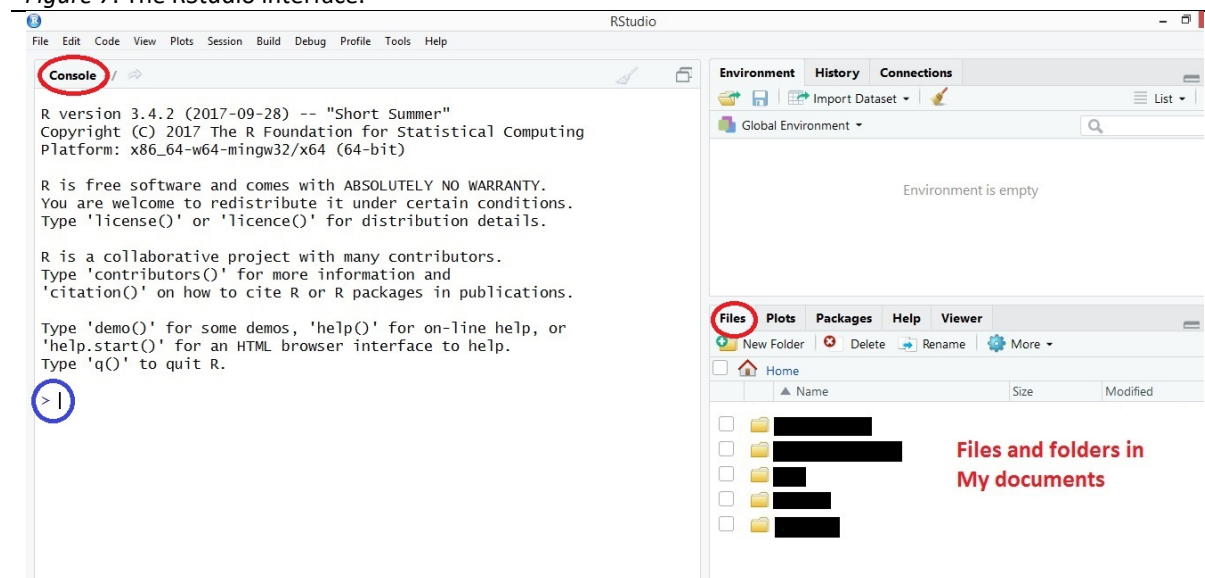
### 3. Getting started with RStudio.

In general, we use RStudio instead of R because of the friendly interface of the former.

To boot up RStudio, we click on the RStudio icon: 

By opening RStudio, we should see something like the Figure 7:

Figure 7. The RStudio interface.



By default, the *Console* is in the left side. After a brief informative text, the system prompt appears (“>”). Can you see the flashing cursor? Here is where we can write the commands or instructions to be run (by R). To execute an instruction in the *Console* and get the result we press *Enter*.

#### 3.1. How to create a script.

Working in the *Console* is very limited since in the *Console* the instructions are generally written and executed one by one. For this reason, it is usual working with **scripts** or instructions files. These files have the extension “.R”.

To create a script we select *File > New File > R Script*.

Now, the script is the top left panel and the *Console* is the bottom left panel. By default, the name of the script is “*Untitled1*” (see Figure 8).


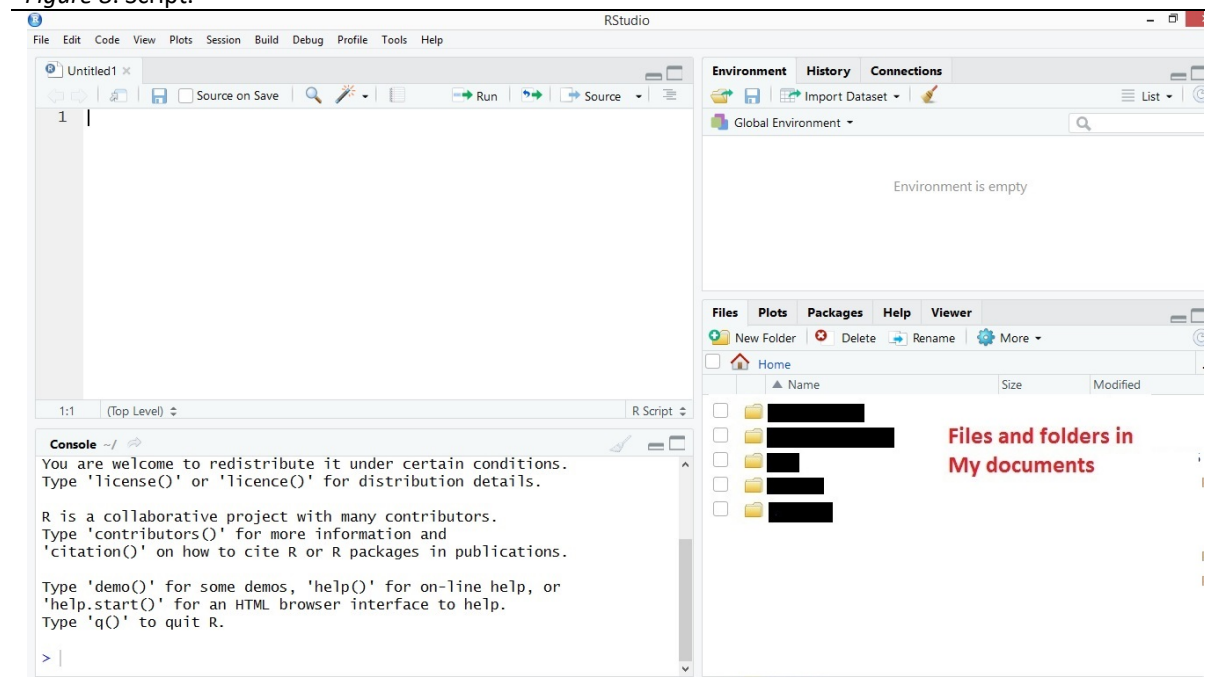
In a script we can write the instructions line by line. The instructions can be run one by one or we can select all (or some) of them and run the selection. We click on  to run them.

Figure 8. Script.



## 4. How to create and work with projects in RStudio.

Once we boot up RStudio, we should set the working directory to indicate R and RStudio where our data, scripts, etc. are. However, in our opinion, the best option is to create a project in RStudio. Creating a project, all the files (data, scripts, etc.) and directories inside the project are directly linked to the project. That is to say, all the work on a project will be self-contained in the directory of the project. In this way, we can easily share the project with someone else or copy-paste the project in another computer, etc.

### Important note:

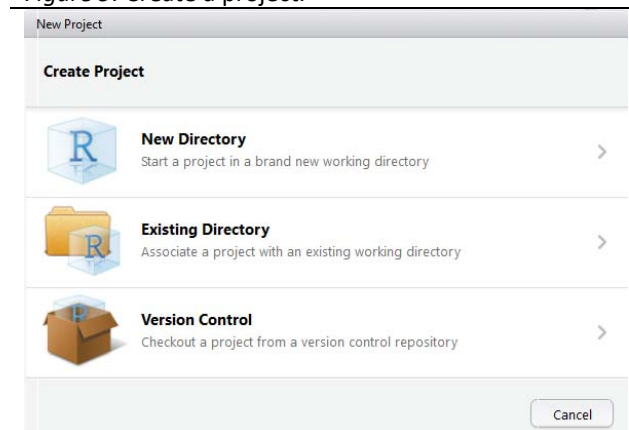
We create a project for each activity/work (for example, for each article/paper).

To organize a project properly, we recommend to create specific directories for raw data, results, text documents, etc.

### 4.1. Creating a project.

To create a project in RStudio, we select *File > New project...* We should see something similar as the Figure below:

Figure 9. Create a project.



To create a project in a new directory, we click on the *New Directory* button. Then, we select the type of project. In our case, *New Project*. Now, we name the directory (or folder) that we are creating. It will also be the name of the project. Finally, we click on *Create Project*. Following this process, a folder in Documents will be created, and inside it we will find the file: "*folder\_name.Rproj*".

To create a project in an existing directory, we click on the *Existing Directory* button, and then we select the directory or folder by clicking on the *Browse...* button. Finally, we click on *Create Project*.

## 4.2. Opening a project.

To open a Project we double click on the file with extension "*.Rproj*". We can also open the project from the RStudio menu: *File > Open Project...*

Advantage of projects: any file we save while we are working in a project will be saved inside the project's directory.

## 4.3. Additional information.

Here you have two readings about projects in RStudio. The second reading includes information about how to create projects.

- <https://www.r-bloggers.com/managing-projects-using-rstudio/>
- [https://www.ssc.wisc.edu/sscc/pubs/RFR/RFR\\_Projects.html#projects](https://www.ssc.wisc.edu/sscc/pubs/RFR/RFR_Projects.html#projects)

## 4.4. How to create a project.

Step 1. Open RStudio

Step 2. Select *File > New Project*

Step 3. Select *New Directory*

Step 4. Project Type: select *New Project*

Step 5. Directory name: Paper\_1

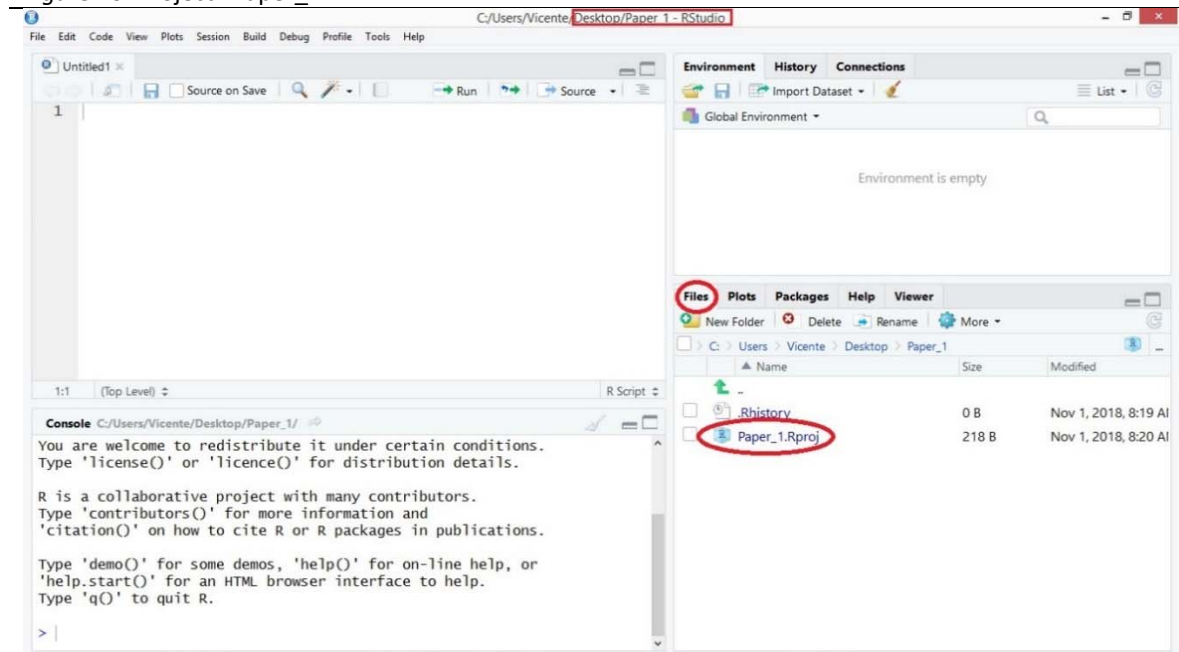
Step 6. Create project as subdirectory of: by using the *Browse*, select the route where the project folder will be created. For example, select the Desktop folder.

Step 7. Click on *Create Project*

**Result:** The “*Paper\_1*” folder will be created in the Desktop folder. This folder will have two files (see Figure 10).

- **Paper\_1** (type: R Project)
- **.Rhistory** (type: RHISTORY file)

Figure 10. Project “*Paper\_1*”.



#### Important note:

The working session with a project will be always opened double-clicking on the file with extension “.*Rproj*” or

*File > Open Project* and selecting the project

Working simultaneously with several projects is not recommended.

## 5. How to install and load dear.

Open RStudio or the project: “*Paper\_1*”.

Select *File > New File > R script*

### 5.1. Installing dear.

To install **dear**, we write into the script:

```
install.packages("dear")
```

and we run the instruction. For that, we click on the *Run* button:  *Run* (see Figure 11).

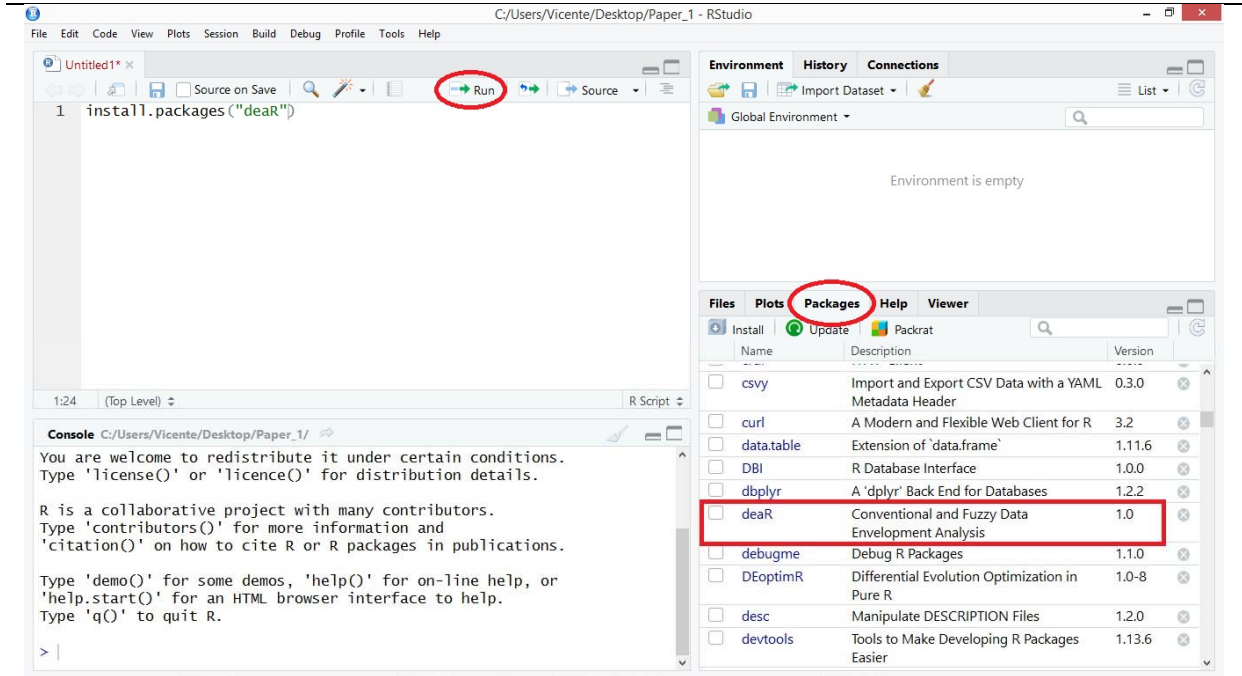
By clicking on the *Packages* menu in the bottom right window, a list of all installed packages will be shown. Now, **dear** package should be listed.



**Important note:**

We have to install **deaR** only once. Then, we can get the new **deaR** versions updating the package.

Figure 11. Install **deaR**.

**5.2. Loading deaR.**

Once we have installed **deaR**, or any other R package, to use it we have to load it first (see Figure 12). For that, we write into the script:

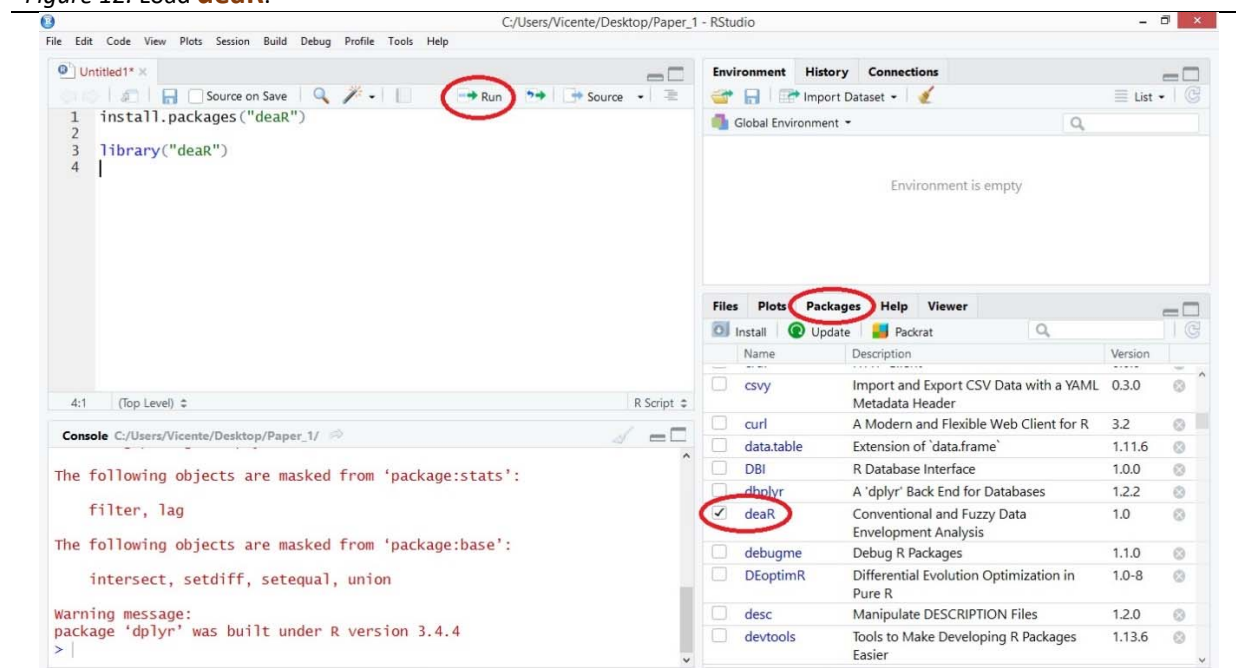
```
library("deaR")
```

and we run the instruction clicking on the Run button: 

**Important note:**

It is necessary to load **deaR** package in each working session.

Figure 12. Load dearR.



## 6. Save the script and close the working session.

To save the script, we select *File > Save as...*

Now, we name the file (for example: *session\_1*) and click on the *Save* button. By default, the script will be saved in the project folder: "*Paper\_1*". Remember, this is an advantage of working with projects.

Notice that the name of the R file (script) is "*session\_1.R*" instead of "*Untitled1*" (see Figure 13). This file is shown in the list of files that are part of the project *Paper\_1*.

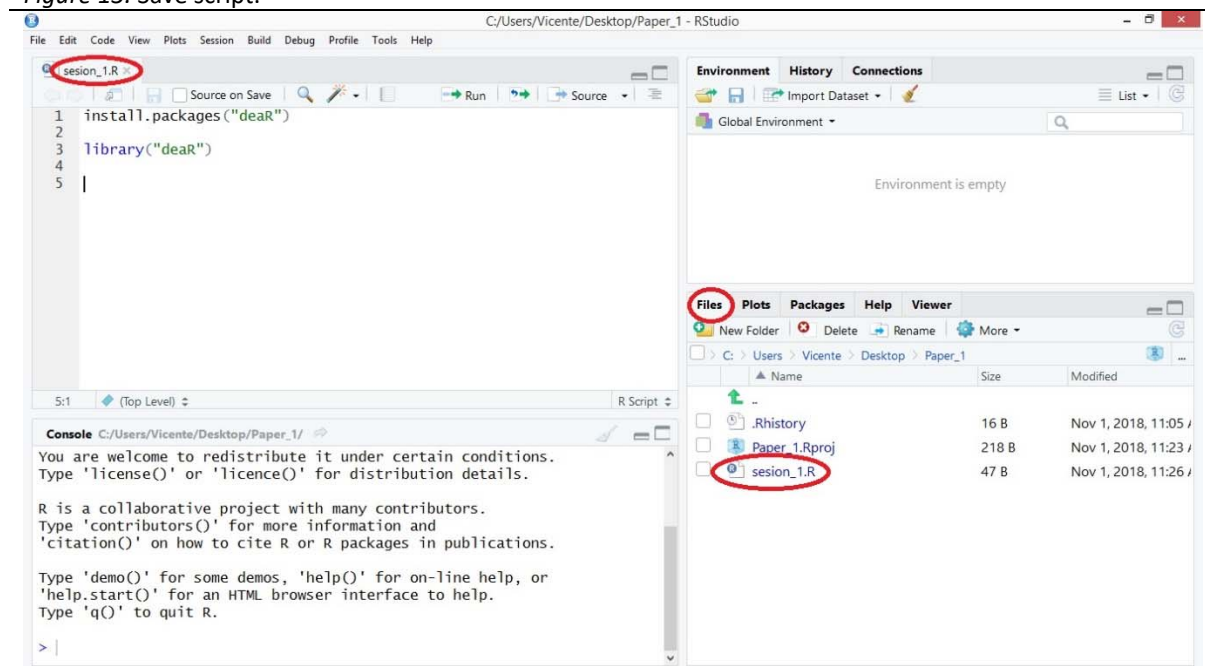
Note: The content of this first script is completely irrelevant up to now. What is important is to remember that we can save a set of instructions in a script.

### Important note:

Give names to scripts that reflect their content.



Figure 13. Save script.



If we want to close the project but not Rstudio: *File > Close project*

If we want to close the project and Rstudio: *File > Quit Session*


At this point, **we save "sesion\_1.R", close the Project "Paper\_1" and quit RStudio.**

## 7. Data Envelopment Analysis with dear.

We start by opening the project "Paper\_1". For that, we double click on the file "Paper\_1.Rproj". RStudio and the project will open. Remember, you can also open RStudio first and then the project (*File > Open Project*).

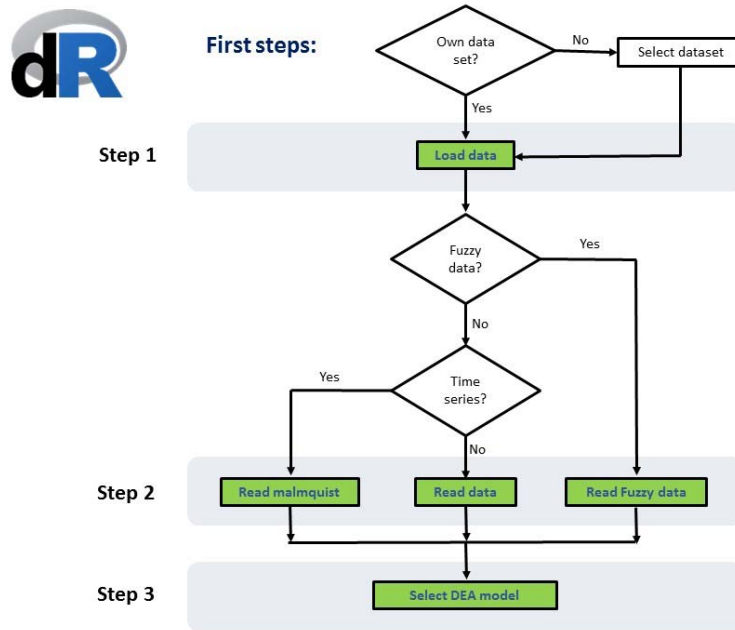
Next, we create a new script (*File > New File > R Script*). In this new script we write the instruction to load **dear**:

```
library("dear")
```

(Note: we execute the instruction by clicking on the *Run* button ).

In the flow chart below (see Figure 14), we can see the steps to perform any DEA (Data Envelopment Analysis) with **dear**.

Figure 14. Steps to use dear.



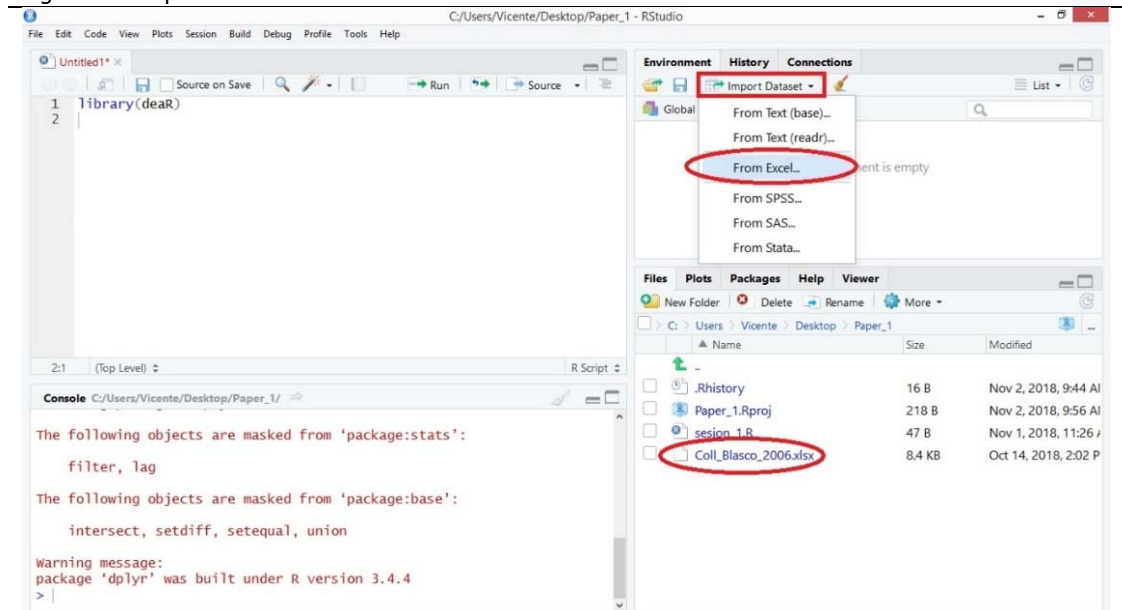
## 7.1. Importing data into R.

The first step is to import the data we are going to use to perform a DEA analysis. The non-R user can use the *Import dataset* option. Let's see how to import data with Example 1.

**Example 1.** Import data into R data from an Excel file:

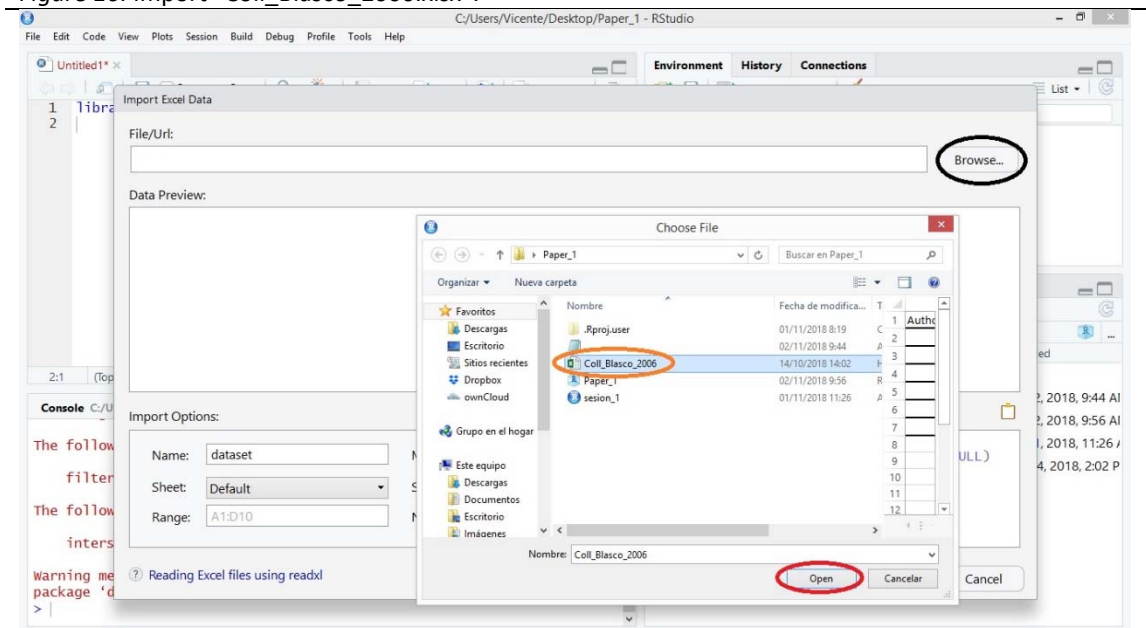
1. Download the data of the example from [www.uv.es/vcoll/Coll\\_Blasco\\_2006.xlsx](http://www.uv.es/vcoll/Coll_Blasco_2006.xlsx) and save the file into the folder of the project "Paper\_1".
2. From RStudio *Environment* menu in the top left window, we select the option: *Import Dataset < From Excel* (see Figure 15).

Figure 15. Import data from Excel.



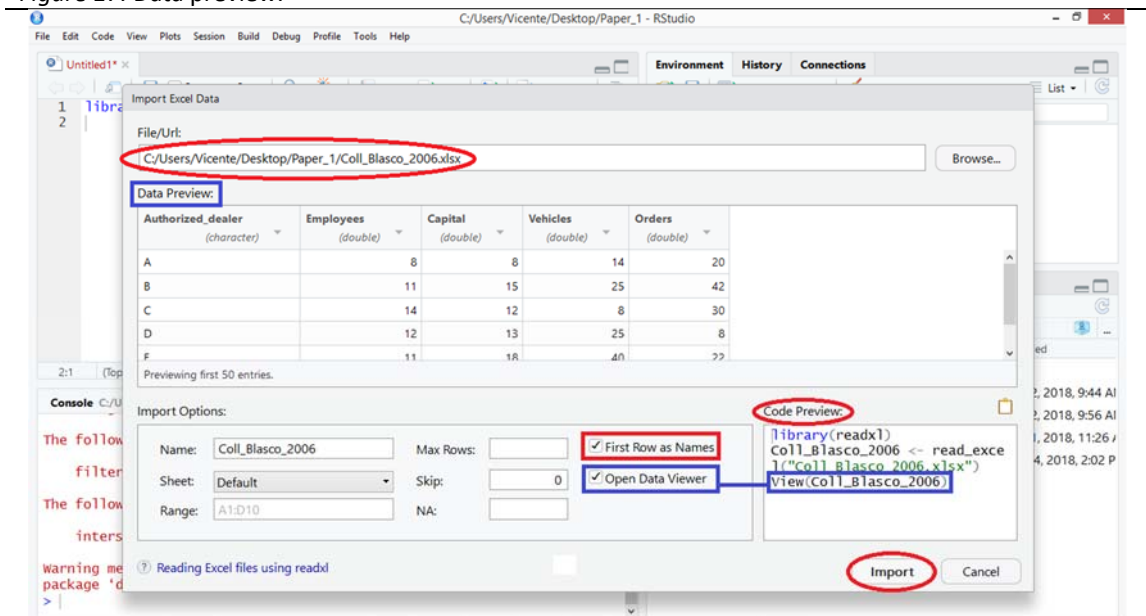
3. The window *Import Excel Data* opens. We click on the *Browse* button to choose the Excel file (“*Coll\_Blasco\_2006.xlsx*”)<sup>2</sup> and then on the *Open* button (see Figure 16).

Figure 16. Import “*Coll\_Blasco\_2006.xlsx*”.



4. Now, we can previsualize “*Coll\_Blasco\_2006.xlsx*” and the import data options used (see left side in Figure 17). As you can see in the Figure below, the R code used to import the data is shown in the bottom right side of this window (see Figure 17). Since the *Open Data Viewer* is selected<sup>3</sup>, the dataset will be opened when the import process concludes. For that, we click on the *Import* button.

Figure 17. Data preview.

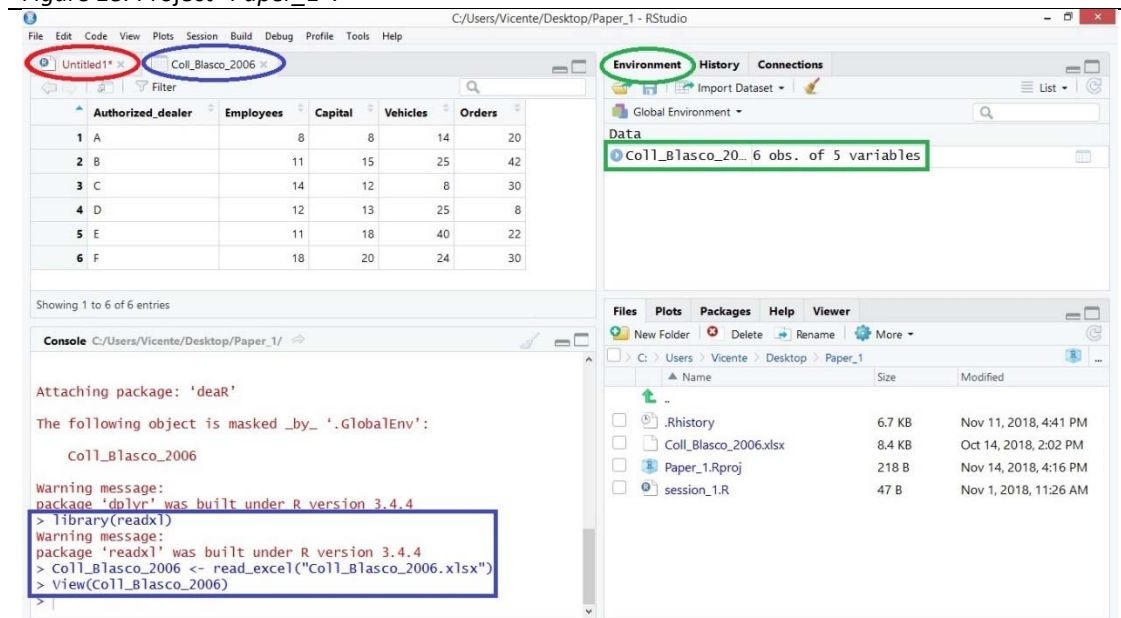


<sup>2</sup> Coll-Serrano, V.; Blasco-Blasco, O. (2006). *Evaluación de la Eficiencia mediante el Análisis Envoltante de Datos. Introducción a los Modelos Básicos*. [www.eumed.net/libros/2006c/197/](http://www.eumed.net/libros/2006c/197/)

<sup>3</sup> It corresponds to the R code: `View(Coll_Blasco_2006)`

5. After clicking on the *Import* button, we come back to the main window of the project “*Paper\_1*”. The imported data are shown in a new sheet (it has the same name as the dataset) (see Figure 18). Besides, in the *Environment* menu (top right panel) appears a list of all objects that we are creating in R. Right now, we only have one object: “*Coll\_Blasco\_2006*”. Actually, this object is a dataframe with 6 observations and 5 variables.

Figure 18. Project “*Paper\_1*”.



We can see the R code used to import the data in the *Console* (see Figure 18):

- **library(readxl)** → load the package: “*readxl*”
- **Coll\_Blasco\_2006 <- read\_excel(“Coll\_Blasco\_2006.xlsx”)** → the “*read\_excel*” function (from the *readxl* package) reads the data file and assign (“<- ”) the data to the object “*Coll\_Blasco\_2006*”.

#### Very important:

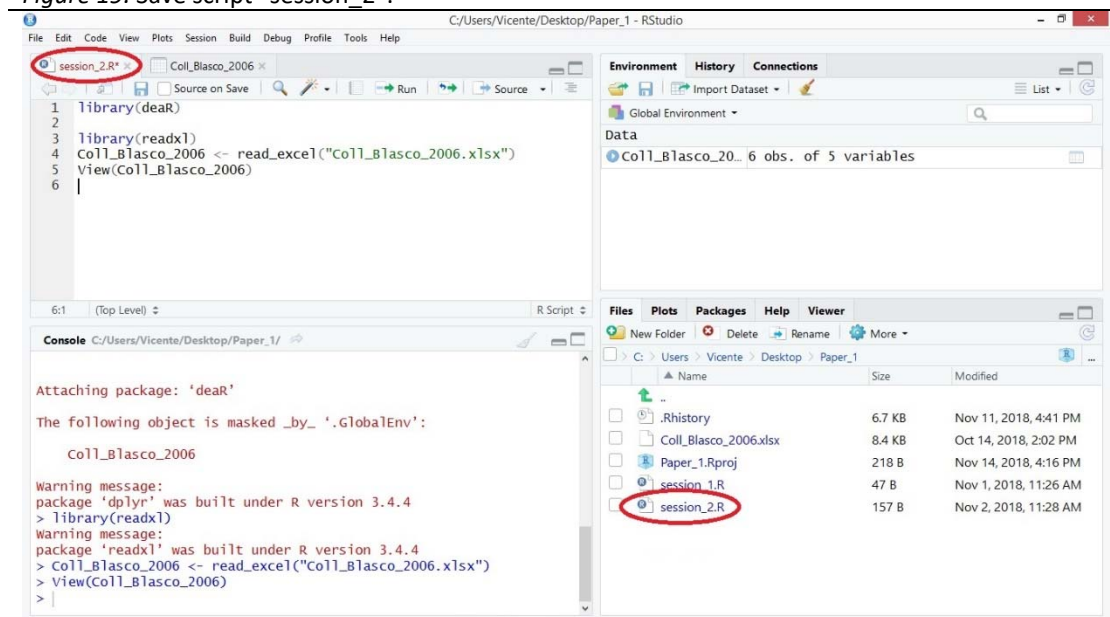
**R terminology:** **A <- B** This means that B (whatever B: a dataset, a function, a matrix, the result of a mathematical operation, etc.) is assigned (<-) to the object A.

- **View(Coll\_Blasco\_2006)** → visualize (View) the object “*Coll\_Blasco\_2006*”.

6. We copy the above code into the script “*Untitled1*” and save it with the name “*session\_2*”.

Our project should show something like that (see Figure 19).

<sup>4</sup> Everything in R is an object. An object can be a dataset, a function, an instruction, a matrix, etc.

Figure 19. Save script “*session\_2*”.

It will not be necessary to repeat steps 1-5 on all subsequent sessions because we have the whole code in the file “*session\_2.R*”. To import the data from “*Coll\_Blasco\_2006.xlsx*” again, we only have to execute the code in “*session\_2.R*”.

**dear** also provides a large collection of datasets. These datasets come from articles already published and they are used to replicate the results of the articles. We think this is a great added value that **dear** offers both researchers and practitioners, since it can be an important help and support for teaching and learning DEA methodology. We can see the structure and source of the data by getting help from **dear**<sup>5</sup>.

Let’s see how to load a dataset in Example 2.

### **Example 2.** Loading datasets available in **dear**.

To see what datasets are available in **dear**, we write

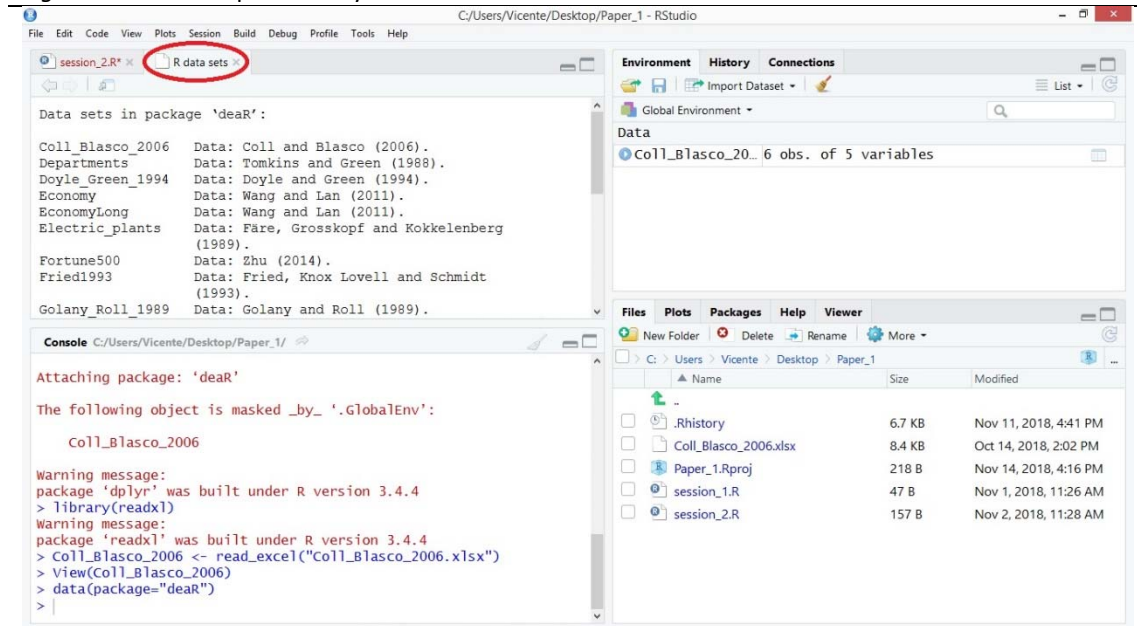
```
data(package="dear")
```

and run the instruction.

<sup>5</sup> Getting help with R: <https://www.r-project.org/help.html>



Figure 20. Datasets provided by dear.



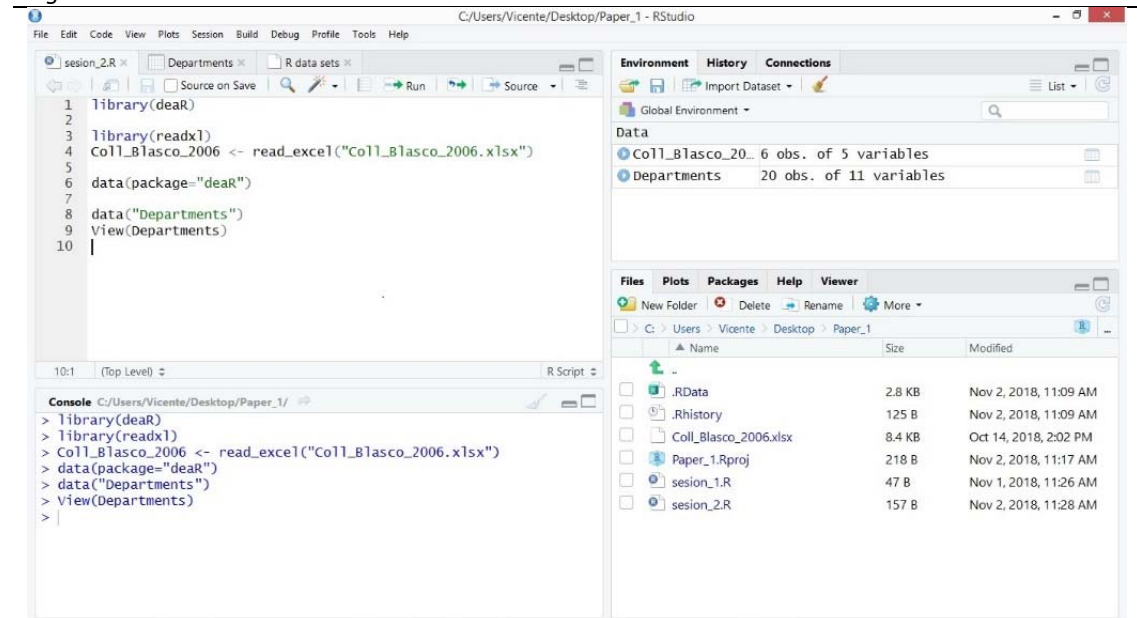
To load a dataset we use the `data()` function. For example, we are going to load the data from Tomkins and Green (1988)<sup>6</sup>. The name of this dataset is "Departments". Therefore, we write into the script "session2.R" (see Figure 21):

```
data("Departments")
```

If we want to visualize the dataset, we write (see Figure 21):

```
View(Departments)
```

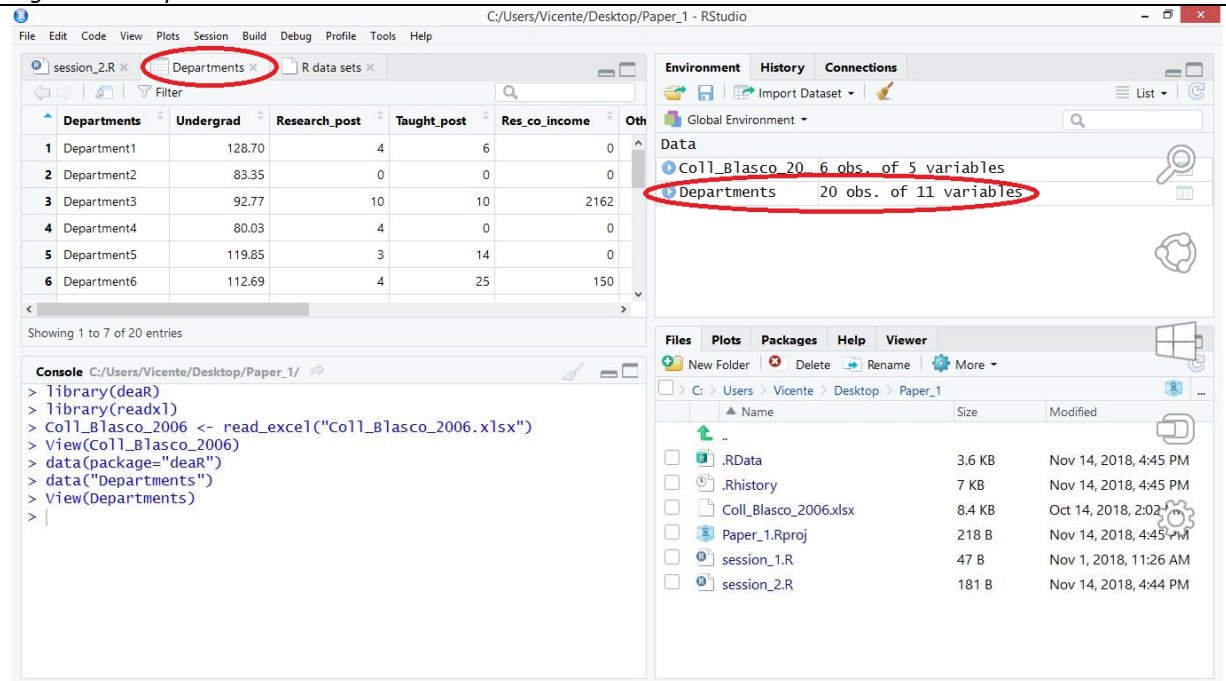
Figure 21. Load and view data.



We execute the instructions by clicking on  (see Figure 22).

<sup>6</sup> Tomkins, C.; Green, R. (1988). "An Experiment in the Use of Data Envelopment Analysis for Evaluating the Efficiency of UK University Departments of Accounting", Financial Accountability and Management, 4(2), 147-164. <https://doi.org/10.1111/j.1468-0408.1988.tb00296.x>

Figure 22. "Departments".



Notice that we now have two objects: "Coll\_Blasco\_2006" and "Departments". "Departments" is a dataframe consisting of 20 observations (DMUs) and 11 variables.

Now, we save "session\_2.R" and close the project "Paper\_1". Quit RStudio.

## 7.2. Adapting the data to deaR

Once the data have been loaded, the next step is to adapt them to the format that **deaR** uses to read them. **deaR** has three different data reading functions. Each reading function responds to a typology of DEA model. These reading functions are:

**Read\_data()**: if we are going to run a conventional DEA model.

**Read\_malmquist()**: if we want to perform the Malmquist productivity index.

**Read\_data\_fuzzy()**: if we want to run a DEA model with uncertain data (fuzzy DEA).

### 7.2.1. Getting help from deaR<sup>7</sup>.

We can learn how to use a specific function by using the help from **deaR**. In the help for **deaR** we can read about the arguments of a function or examples about how to use it. Let's see how to use the help function with Example 3.

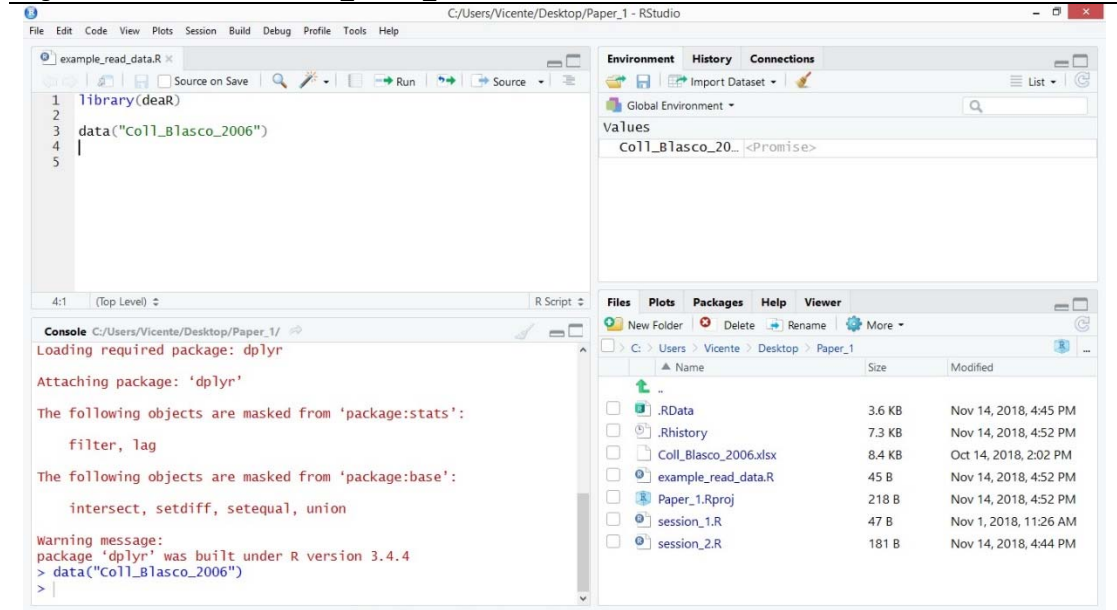
#### Example 3. Getting help for deaR.

1. We open the project "Paper\_1" and create a new script. We name the script: "example\_read\_data"
2. We load **deaR**
3. We load the dataset provided by **deaR**: "Coll\_Blasco\_2006" (see Figure 23).

<sup>7</sup> Getting help with R: <https://www.r-project.org/help.html>



Figure 23. Load dataset: "Coll\_Blasco\_2006"

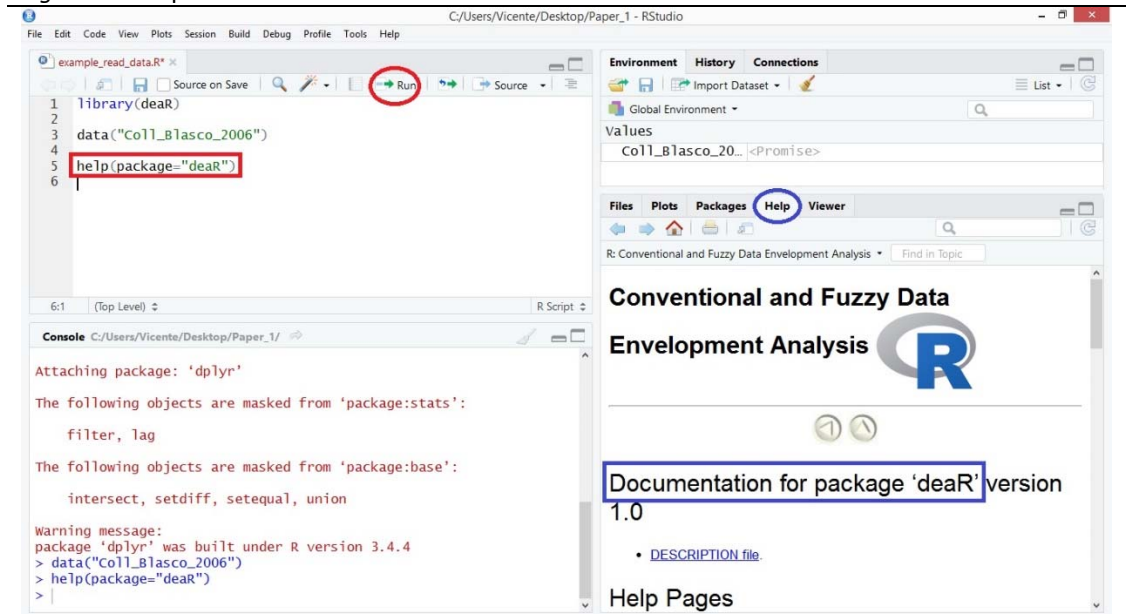


To access the documentation for a function of the package **deaR** we use the **help()** function. We write:

```
help(package="deaR")
```

In the *Help* menu (see bottom right window) a list of all functions and datasets in **deaR** will appear. It is the documentation for the package (see Figure 24).

Figure 24. Help for deaR.



### 7.2.2. `read_data()` function:

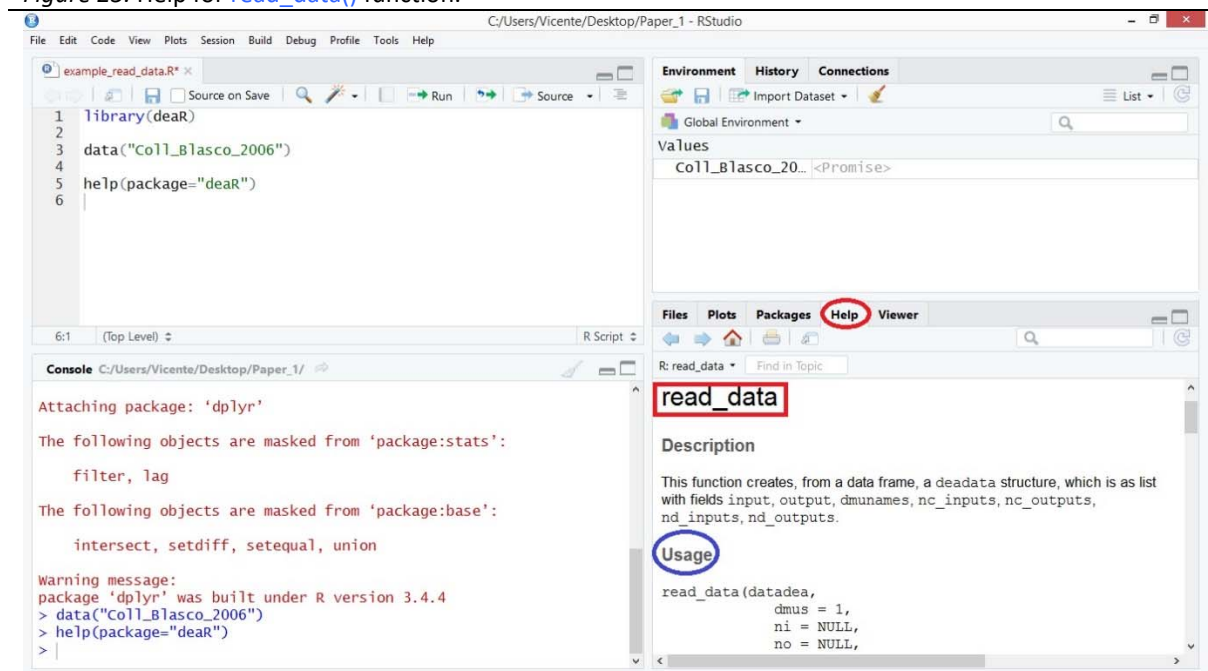
Once we have accessed the documentation (help) for **dearR**, if we clic on the “`read_data`” link we will get the specific help about this function. We can also get the same result if we write:

**`help(read_data)`** (or **`?read_data`**)

into the script and run the instruction.

All the arguments of the `read_data()` function are listed in the *Usage* section, and in the *Arguments* section we can read a brief explanation of them (see Figure 25).

Figure 25. Help for `read_data()` function.



The `read_data()` function takes the following arguments:

- *datadea*: It is the dataset (It must be a dataframe).
- *dmus*: Number of the column where the DMUs are. By default, **dearR** considers DMUs are in the first column of the dataset.
- *ni*: It is the number of inputs.
- *no*: It is the number of outputs.
- *inputs*: We can indicate the numbers of the columns where the inputs are instead of indicating the number of inputs.
- *Outputs*: We can indicate the numbers of the columns where the outputs are instead of indicating the number of outputs.
- *nc\_inputs*: If we have non-controllable inputs, we can indicate which are them.
- *nc\_outputs*: If we have non-controllable outputs, we can indicate which are them.
- *nd\_inputs*: If we have non-discretionary inputs, we can indicate which are them.
- *nd\_outputs*: If we have non-discretionary outputs, we can indicate which are them.
- *ud\_inputs*: If we have undesirable (bad) inputs, we can indicate which are them.

- *ud\_outputs*: If we have undesirable (bad) outputs, we can indicate which are them.

Example 4 shows how to use the `read_data()` function. The documentation for **dear** provides examples for the all functions in the package.

**Example 4.** Using the `read_data()` function.

Currently, we have loaded the dataset “*Coll\_Blasco\_2006*”. This dataset has 6 DMUs (column 1), 2 inputs (columns 2 and 3) and 2 outputs (columns 4 and 5).

Suppose that we want to measure the efficiency of these DMUs using the BCC DEA model.

As BCC model is a conventional DEA model (it is not a fuzzy DEA model), we must use the `read_data()` function to adapt the original dataset to the format used by **dear**.

We write the following text into the script “*example\_read\_data*” (see Figure 26):

```
data_example <- read_data(Coll_Blasco_2006, ni=2, no=2)
```

**Very important:** Reading and understanding the instruction.

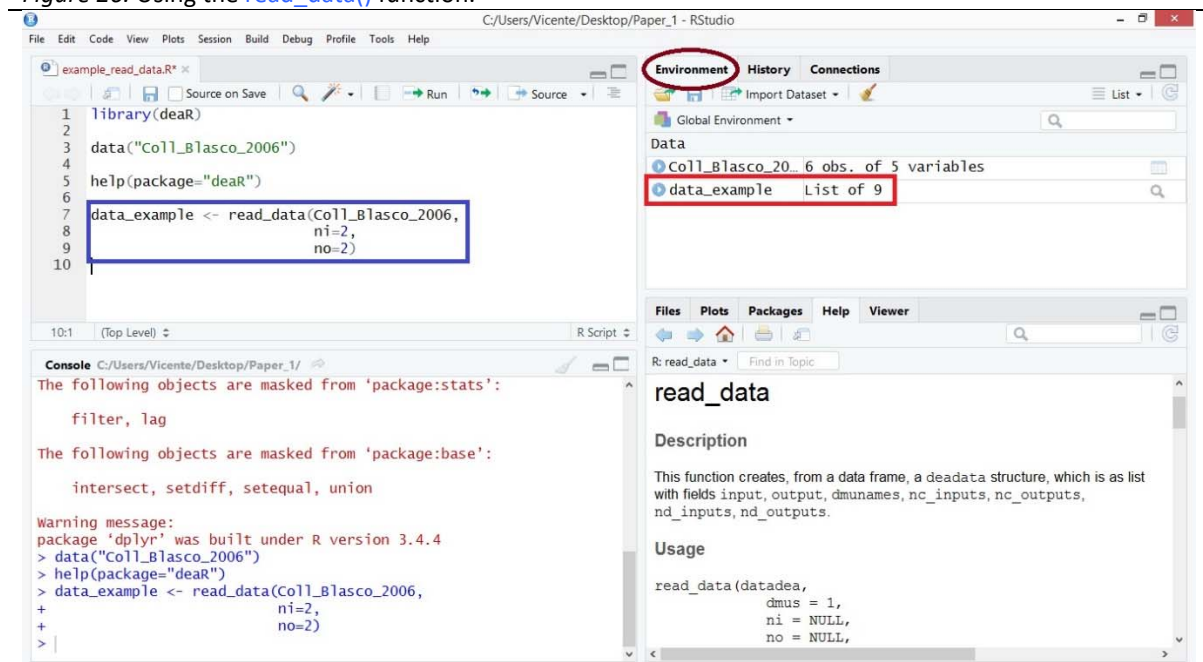
The expression on the right side of the assignment symbol (<-) means: read the data (`read_data`) in *Coll\_Blasco\_2006*; it has 2 inputs (`ni=2`) and 2 outputs (`no=2`). DMUs are in column 1 (`dmus=1`). This argument doesn't appear since it is the default value.

Then, the result of this function is assigned (<-) to the object *data\_example*.

We execute the instruction (  ).

Notice that now the objects “*Coll\_Blasco\_2006*” and “*data\_example*” appear listed in the *Environment* menu. Notice also that “*data\_example*” is a list of 9 components. (see Figure 26).

Figure 26. Using the `read_data()` function.



Now, the data are ready to be used with a conventional DEA model (see section 7.3.). For that, we should use the object we have created: `"data_example"`.

Recommendation: Practice the function `read_data()` with the examples in the documentation.

---

Save `"example_read_data.R"`.

### 7.2.3. `read_malmquist()` function.

If we have time series data and we want to measure the Malmquist productivity index, we must use the `read_malmquist()` function first to adapt them to the format that **deaR** reads.

**deaR** admits the time series data in two different formats:

- **Wide format:** DMUs by column. Inputs and outputs of different periods of time by column. For example, see the dataset `"Economy"`<sup>8</sup> in **deaR** (Figure 27).
- **Long format:** Time by column. DMUs, inputs and outputs by column, but grouped by time. For example, see the dataset `"EconomyLong"` in **deaR** (Figure 28).

In the following example we try to show the above explanation.

---

#### **Example 5.** Wide and Long data formats.

Now, we are going to show these different data formats. For that:

1. We create a new script. Name it: `"example_read_malmquist"`.

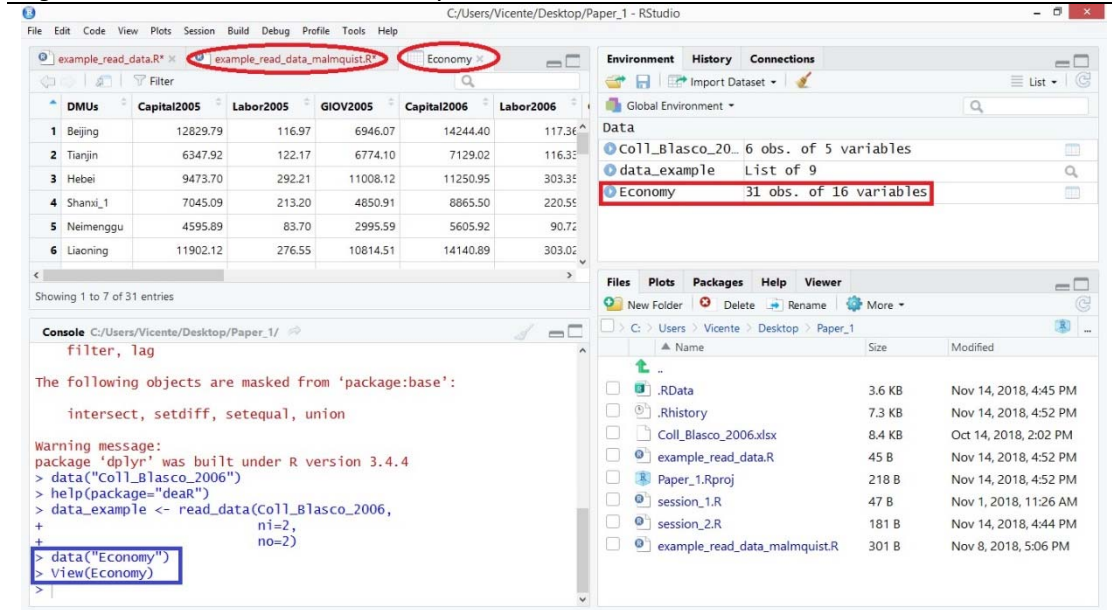
Note: If we have closed the working session, we open the project `"Paper_1"` and then create the new script. We also have to load **deaR**.

2. We load the dataset in **deaR**: `"Economy"` (see Figure 27).
3. Visualize `"Economy"`. This object (dataset) has 31 DMUs with 2 inputs (Capital and Labor) and 1 output (GIOV) for 5 years (from 2005 to 2009). `"Economy"` is a dataset in wide format.

---

<sup>8</sup> Wang, Y.; Lan, Y. (2011). "Measuring Malmquist Productivity Index: A New Approach Based on Double Frontiers Data Envelopment Analysis". *Mathematical and Computer Modelling*, 54, 2760-2771. <https://doi.org/10.1016/j.mcm.2011.06.064>

Figure 27. Wide data format: "Economy".



The screenshot shows the RStudio interface with the 'Economy' dataset loaded. The Environment pane on the right shows 'Economy' with 31 observations and 16 variables. The console shows the following code and output:

```

filter, lag

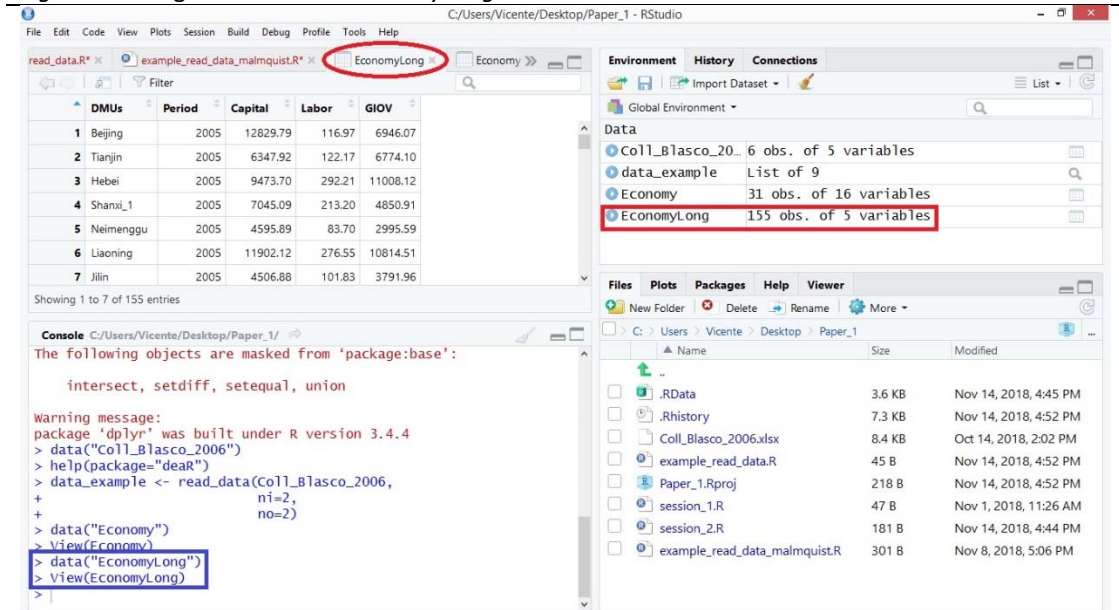
The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

Warning message:
package 'dplyr' was built under R version 3.4.4
> data("Coll_Blasco_2006")
> help(package="dear")
> data_example <- read_data(Coll_Blasco_2006,
+                             ni=2,
+                             no=2)
> data("Economy")
> View(Economy)

```

- Now, load the dataset "EconomyLong".
- We visualize "EconomyLong" (see Figure 28). This new R object (actually, it is a dataset) has 155 observations (31 DMUs x 5 years), with 2 inputs (Capital and Labor) and 1 output (GIOV). The *Period* column refers to the time period 2005-2009.

Figure 28. Long data format: "EconomyLong".



The screenshot shows the RStudio interface with the 'EconomyLong' dataset loaded. The Environment pane on the right shows 'EconomyLong' with 155 observations and 5 variables. The console shows the following code and output:

```

read_data.R* x example_read_data_malmquist.R* EconomyLong Economy

The following objects are masked from 'package:base':
  intersect, setdiff, setequal, union

Warning message:
package 'dplyr' was built under R version 3.4.4
> data("Coll_Blasco_2006")
> help(package="dear")
> data_example <- read_data(Coll_Blasco_2006,
+                             ni=2,
+                             no=2)
> data("Economy")
> View(Economy)
> data("EconomyLong")
> View(EconomyLong)

```

Note: Notice that all objects we are creating during the working session are listed in the *Environment* menu (top right window).



The `read_malmquist()` function takes the following arguments<sup>9</sup>:

- *datadea*: It is the dataset (It must be a dataframe).
- *nper*: It is the number of years (with data in wide format).
- *percol*: It is the number of the column where the variable time is (with data in long format).
- *arrangement*: “horizontal” with data in wide format or “vertical” with data in long format.
- *dmus*: Number of the column where the DMUs are. By default, **deaR** considers DMUs are in the first column of the dataset.
- *ni*: It is the number of inputs.
- *no*: It is the number of outputs.
- *inputs*: We can indicate the numbers of the columns where the inputs are instead of indicating the number of inputs.
- *Outputs*: We can indicate the numbers of the columns where the outputs are instead of indicating the number of outputs.
- *nc\_inputs*: If we have non-controllable inputs, we can indicate which are them.
- *nc\_outputs*: If we have non-controllable outputs, we can indicate which are them.
- *nd\_inputs*: If we have non-discretionary inputs, we can indicate which are them.
- *nd\_outputs*: If we have non-discretionary outputs, we can indicate which are them.
- *ud\_inputs*: If we have undesirable (bad) inputs, we can indicate which are them.
- *ud\_outputs*: If we have undesirable (bad) outputs, we can indicate which are them.

The current version of **deaR** cannot manage undesirable inputs/outputs to calculate the Malmquist productivity index. This feature will be incorporated in a future version.

In Examples 6 and 7 we can see how to use the `read_malmquist()` function with wide and long data respectively.

---

**Example 6.** `read_malmquist()` function with wide data format.

For this example, we are going to use the dataset “*Economy*”. This dataset is already loaded in the current working session.

To adapt “*Economy*” to the reading format used by **deaR**, we write into the script (“*example\_read\_malmquist*”) the following text:

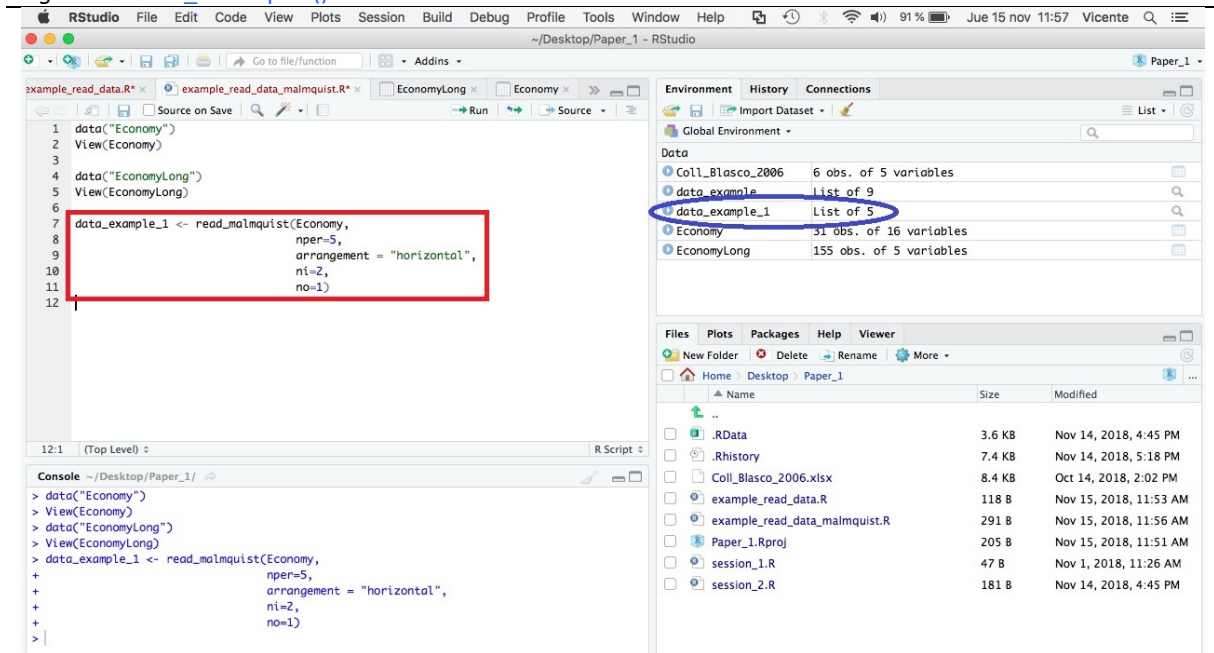
```
data_example_1 <- read_malmquist(Economy,
                                nper=5,
                                arrangement="horizontal",
                                ni=2,
                                no=1)
```

---

<sup>9</sup> See `help(read_malmquist)`.

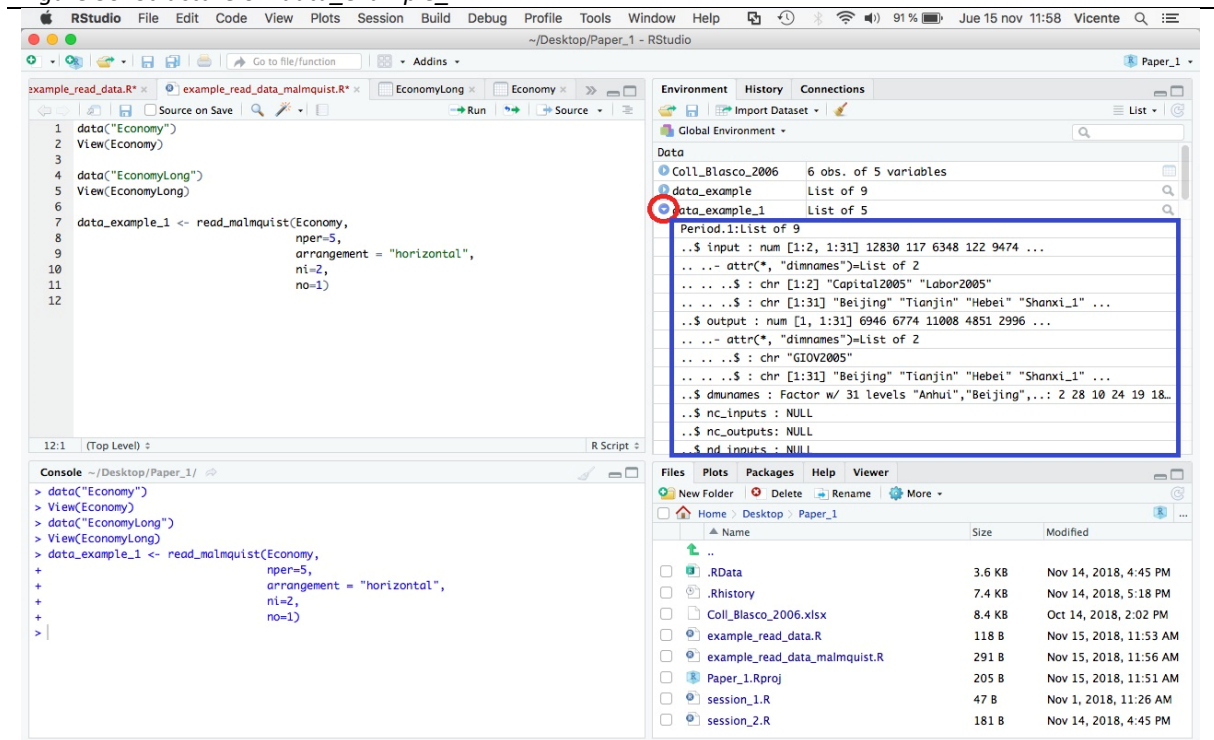
By running the instruction, a new object is created (“*data\_example\_1*”) and listed in the *Environment* (see Figure 29).

Figure 29. `read_malmquist()` function with wide data format.



As we can see, “*data\_example\_1*” is a list of 5 components. If we click on the icon next to the name of the object, we can see the structure of this object (see Figure 30).

Figure 30. Structure of “*data\_example\_1*”.

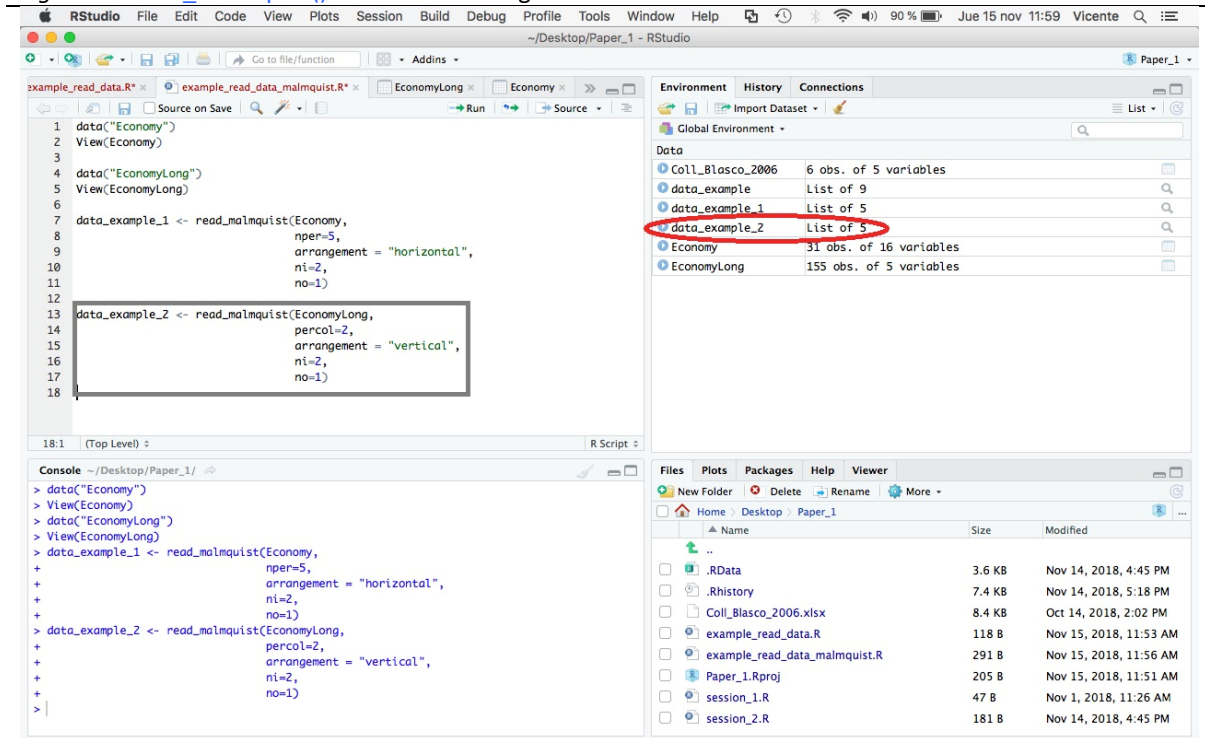




### Example 7. `read_malmquist()` function with long data format.

Now, we are going to use the dataset “*EconomyLong*”. The Figure 31 shows the instruction to adapt the dataset to the reading format in **dear**.

Figure 31. `read_malmquist()` function with long data format.



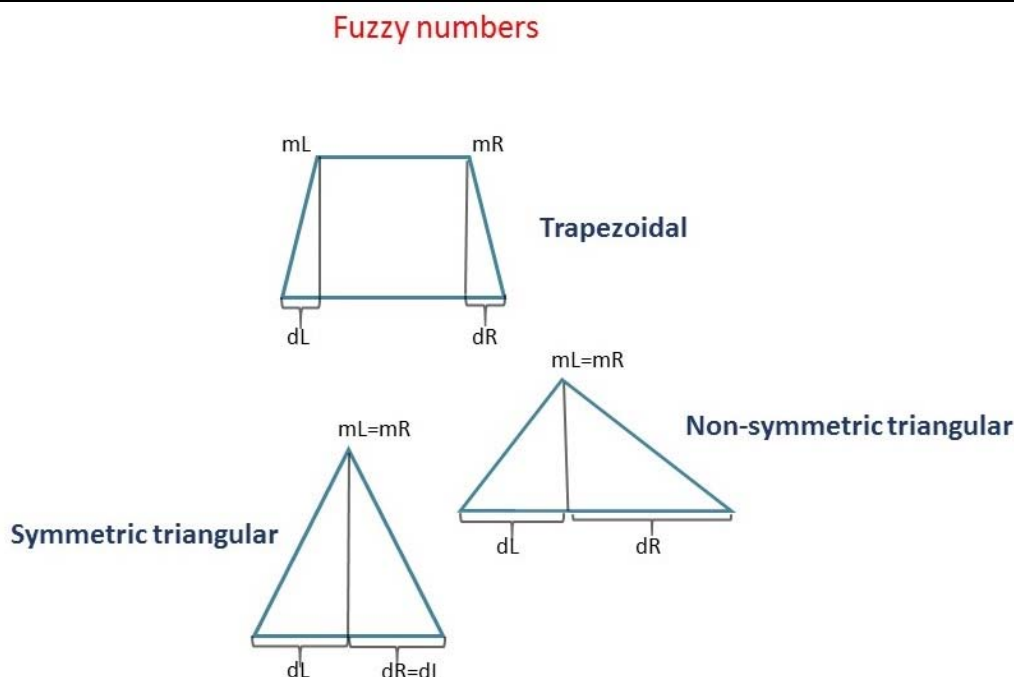
Save “`example_read_malmquist.R`”.

#### 7.2.4. `read_data_fuzzy()` function.

If we are working with uncertain data and we want to measure the efficiency with a fuzzy DEA model so that **dear** can read the data, we have to use the `read_data_fuzzy()` function to adapt them first.

**dear** can manage trapezoidal, symmetric triangular, and non-symmetric triangular fuzzy numbers (see Figure 32).

Figure 32. Fuzzy numbers.



The `read_data_fuzzy()` function takes the following arguments<sup>10</sup>:

- *datadea*: It is the dataset (It must be a dataframe).
- *dmus*: Number of the column where the DMUs are. By default, **deaR** considers DMUs are in the first column of the dataset.
- *Inputs.mL*: Number of the columns where the  $mL$  values of the inputs are.
- *Inputs.mR*: Number of the columns where the  $mR$  values of the inputs are.
- *Inputs.dL*: Number of the columns where the  $dL$  values of the inputs are.
- *Inputs.dR*: Number of the columns where the  $dR$  values of the inputs are.
- *Outputs.mL*: Number of the columns where the  $mL$  values of the outputs are.
- *Outputs.mR*: Number of the columns where the  $mR$  values of the outputs are.
- *Outputs.dL*: Number of the columns where the  $dL$  values of the outputs are.
- *Outputs.dR*: Number of the columns where the  $dR$  values of the outputs are.
- *nc\_inputs*: If we have non-controllable inputs, we can indicate which are them.
- *nc\_outputs*: If we have non-controllable outputs, we can indicate which are them.
- *nd\_inputs*: If we have non-discretionary inputs, we can indicate which are them.
- *nd\_outputs*: If we have non-discretionary outputs, we can indicate which are them.
- *ud\_inputs*: If we have undesirable (bad) inputs, we can indicate which are them.

<sup>10</sup> See [help\(read\\_data\\_fuzzy\)](#).

- *ud\_outputs*: If we have undesirable (bad) outputs, we can indicate which are them.

Currently, **dearR** only considers the undesirable inputs/outputs in the fuzzy DEA models based on the BCC DEA model.

Let's see it in Example 8.

### Example 8. `read_data_fuzzy()` function.

1. We create a new script. Name it: "*example\_read\_data\_fuzzy*".

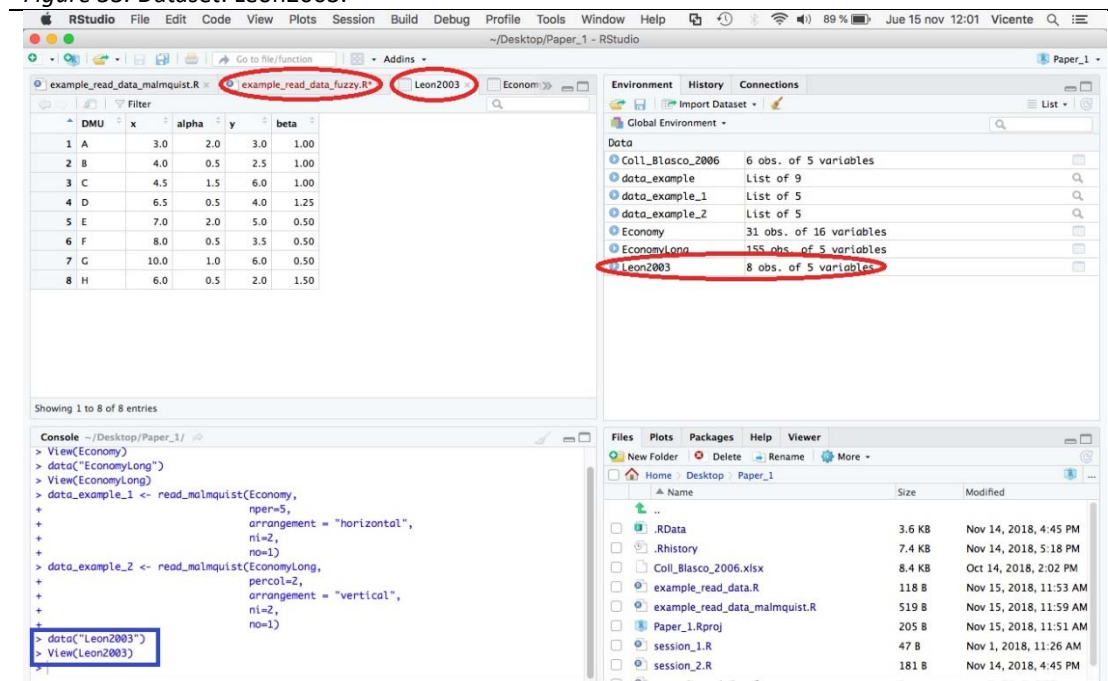
Note: If we have closed the working session, we open the project "*Paper\_1*" and then create the new script. We also have to load **dearR**.

2. We load the dataset "*Leon2003*"<sup>11</sup>:

`data("Leon2003")`.

This dataset (see Figure 33) has 8 DMUs with 1 symmetric triangular fuzzy input (alpha is the spread of the input) and 1 symmetric triangular fuzzy output (beta is the spread of the output).

Figure 33. Dataset: Leon2003.



3. Now, we `read_data_fuzzy()` to adapt the data. For that, we write into the script:

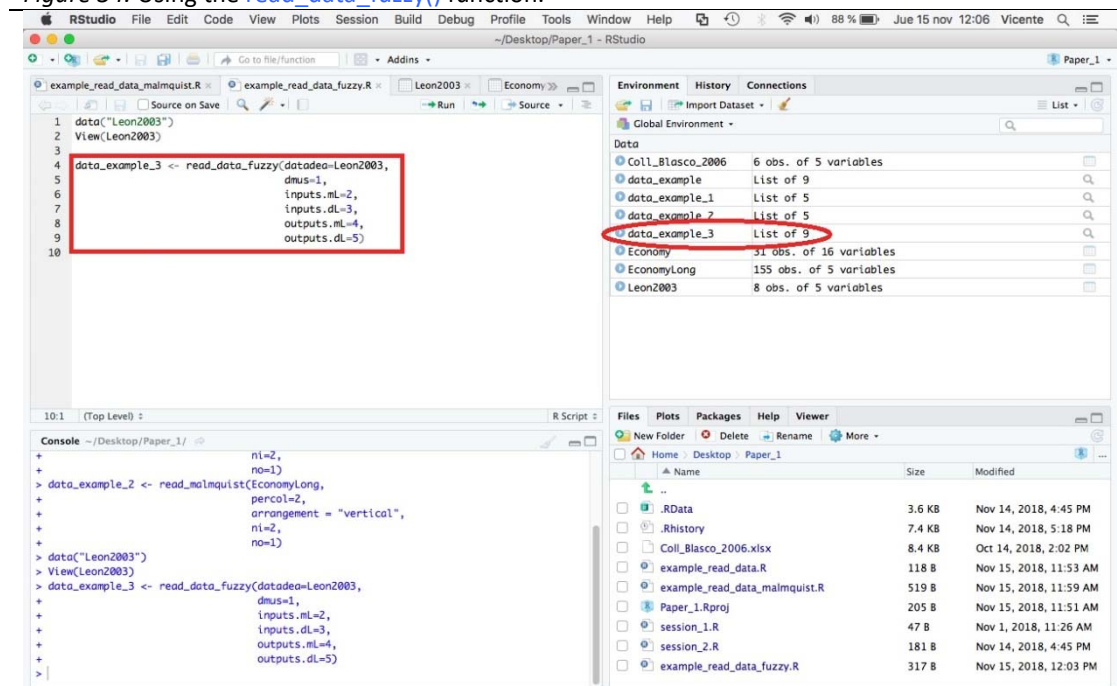
```
data_example_3 <- read_data_fuzzy(Leon2003,
                                  inputs.mL=2,
                                  inputs.dL=3,
                                  outputs.mL=4,
                                  outputs.dL=5)
```

and run it.

<sup>11</sup> León, T.; Liern, V. Ruiz, J.; Sirvent, I. (2003). "A Possibilistic Programming Approach to the Assessment of Efficiency with DEA Models", *Fuzzy Sets and Systems*, 139, 407–419. [https://doi.org/10.1016/S0165-0114\(02\)00608-5](https://doi.org/10.1016/S0165-0114(02)00608-5)

Our working session should look like the screenshot below (Figure 34):

Figure 34. Using the `read_data_fuzzy()` function.



In this example,  $\text{inputs.dL}=\alpha$  and  $\text{outputs.dL}=\beta$ . Since we have symmetric triangular fuzzy numbers:  $\text{inputs.dL}=\text{inputs.dR}$  and  $\text{outputs.dL}=\text{outputs.dR}$

We have created a new R object: `"data_example_3"`. This object will be the one we will use to run a given fuzzy DEA model.

We save `"example_read_data_fuzzy.R"`, close the project and quit RStudio.

### 7.3. Select and run a DEA model.

Once we have adapted the data to the reading format in **deaR**, the next step is to select the DEA model and run it.

In the current version of **deaR** (version 1.0), the following DEA models are available:

#### Conventional DEA models

Basic (radial) models (envelopment and multiplier forms)

Directional distance function model

(Weighted) Additive model

Super-efficiency additive model

Radial Super-efficiency model

(Weighted) Non-radial model

Preference Structure model

(Weighted) Slack-based model

(Weighted) Super-efficiency slack-based model

Conventional DEA models
Cross-efficiency (crs <sup>12</sup> and vrs <sup>13</sup> )
Bootstrapping (Simar and Wilson algorithm)
FDH model
Productivity
Malmquist index
Fuzzy DEA models
Kao and Liu model <sup>14</sup>
Possibilistic model
Guo and Tanaka model
Fuzzy cross-efficiency <sup>15</sup> (only crs)

In the documentation for **deaR** we can find all the details about how to use the different functions and examples putting them into practice.

Although the DEA models available in **deaR** are the listed above, given the flexibility in programming **deaR** (here is an important strength of the package), the user himself can experiment, try and implement variants of these models. For example, if the user defines the weights appropriately in the “*Additive model*”, MPI<sup>16</sup> or RAM<sup>17</sup> models can be obtained. Another example where the user can define weights and obtain different models is from “*Preference Structure Model*” (weighted non radial model). In this case, the user can calculate the “*Cost Efficiency*”, “*Revenue Efficiency*” or “*Profit Efficiency*” models.

Now, we are going to see how to use the `model_basic()` function to perform basic DEA models. For that:

- Open the project “*Paper\_1*” and create a new script. Give it the name: “*example\_basic*”.
- Load **deaR**.
- Go to the help for **deaR**.
- Click on the `model_basic` link (see Figure 35). We can also write into the script:

`help(“model_basic”)`

<sup>12</sup> crs = constant returns-to-scale.

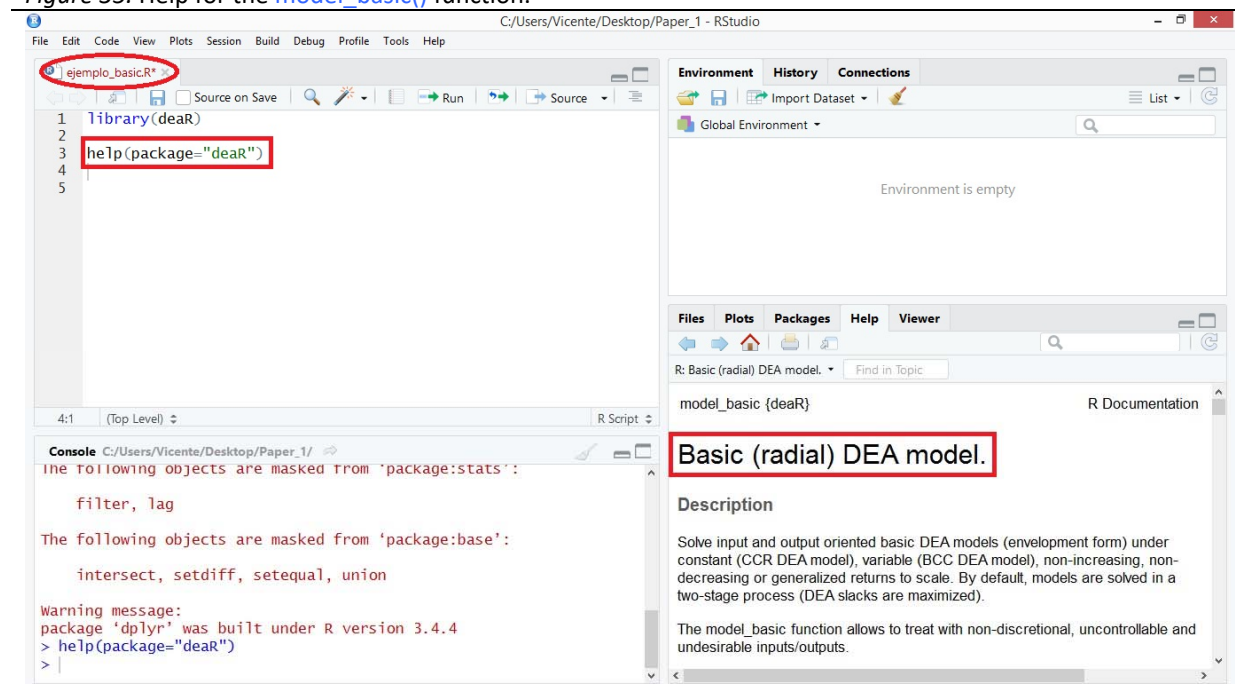
<sup>13</sup> vrs = variable returns-to-scale.

<sup>14</sup> This Fuzzy DEA model has been extended to a several DEA models. See the help for the package: `help(“modelfuzzy_kaoliu”)`.

<sup>15</sup> Based on Guo and Tanaka’s model.

<sup>16</sup> Measure of Inefficiency Proportions (MPI).

<sup>17</sup> Range Adjusted Measure (RAM).

Figure 35. Help for the `model_basic()` function.

The `model_basic()` function takes the following arguments:

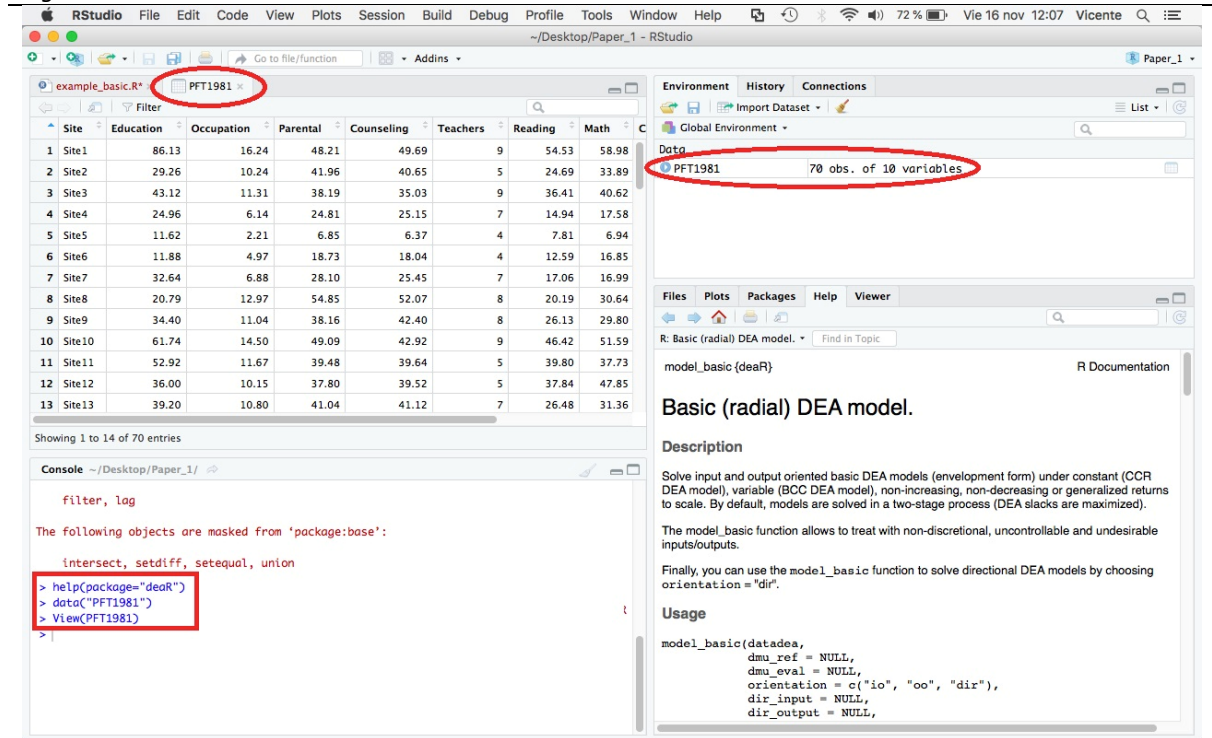
- `datadea`: It is the dataset (It must be a dataframe).
- `dmu_ref`: Selection of a subset of DMUs.
- `dmu_eval`: DMUs to be evaluated from the selected subset.
- `orientation`: Orientation of the DEA model: input-oriented, output-oriented or direccional.
- `dir_input`: Input direction vector in directional models.
- `dir_output`: Output direction vector in directional models.
- `rts`: Returns to scale (constant, variable, non-increasing, non-decreasing, generalized).
- `L`: If we select generalized rts, lower bound.
- `U`: If we select generalized rts, upper bound.
- `Maxslack`: By default, slacks are maximized in a second stage.
- `weight_slack_i`: The user can set weights for the input slacks to be maximized in the second stage.
- `weight_slack_o`: The user can set weights for the output slacks to be maximized in the second stage.
- `vtrans_i`: With undesirable inputs and variable returns-to-scale, the user can set a translation vector. By default, the translation vector is  $(\max + 1)$ .
- `vtrans_o`: With undesirable outputs and variable returns-to-scale, the user can set a translation vector. By default, the translation vector is  $(\max + 1)$ .
- `compute_target`: To calculate the targets for the max slack solution.
- `compute_multiplier`: It returns the multipliers (or weights) of the multiplier form.



- *returnlp*: For each DMU, this argument returns the linear problem of the first stage.

Now, we are going to learn how to use the `model_basic()` function by doing Examples 9 and 10. To follow these examples, we have to load the dataset in **dearR**: “PFT1981”<sup>18</sup>. The dataset has 70 DMUs (school sites) with 5 inputs (*Education, Occupation, Parental, Counseling, Teachers*) and 3 outputs (*Reading, Math, Coopersmith*). The Figure below (Figure 36) shows the instructions that we should write into the script “*example\_basic*” to load the data.

Figure 36. Dataset: PTF1981.



The screenshot shows the RStudio interface with the following components:

- Environment Pane:** Shows the loaded data frame `PFT1981` with 70 observations and 10 variables.
- Console:** Contains the following R commands and their output:
 

```
filter, lag
The following objects are masked from 'package:base':
intersect, setdiff, setequal, union
> help(package="dearR")
> data("PFT1981")
> View(PFT1981)
>
```
- Viewer Pane:** Displays the documentation for the `model_basic(dearR)` function, including a description and usage examples.

### Example 9. `model_basic()` function.

Of 70 DMUs (school sites) in “PFT1981”, there are 49 DMUs in Project Follow Through (PFT) and 21 DMUs in Non-Follow Through (NFT).

In this example, we are going to calculate the efficiency of the 49 DMUs in PFT by using the input-oriented CCR DEA model.

First, we have to use the `read_data()` function to adapt the data to the reading format in **dearR** (see section 7.2.2.). Run the instruction by clicking the **Run** button on

Now, we have to use the `model_basic()` function because we want to measure the efficiency by using the CCR DEA model (a conventional DEA model). But we want to calculate the efficiency of the first 49 DMUs; that is to say, the DMUs participating in PFT. To make this selection, we use the argument: `dmu_ref=1:49`. To evaluate all these units, we use the argument: `dmu_eval=1:49`. Therefore, we have to write into the script “*example\_basic*” the following text:

<sup>18</sup> Charnes, A.; Cooper, W.W.; Rhodes, E. (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through", *Management Science*, 27(6), 668-697. <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.27.6.668>

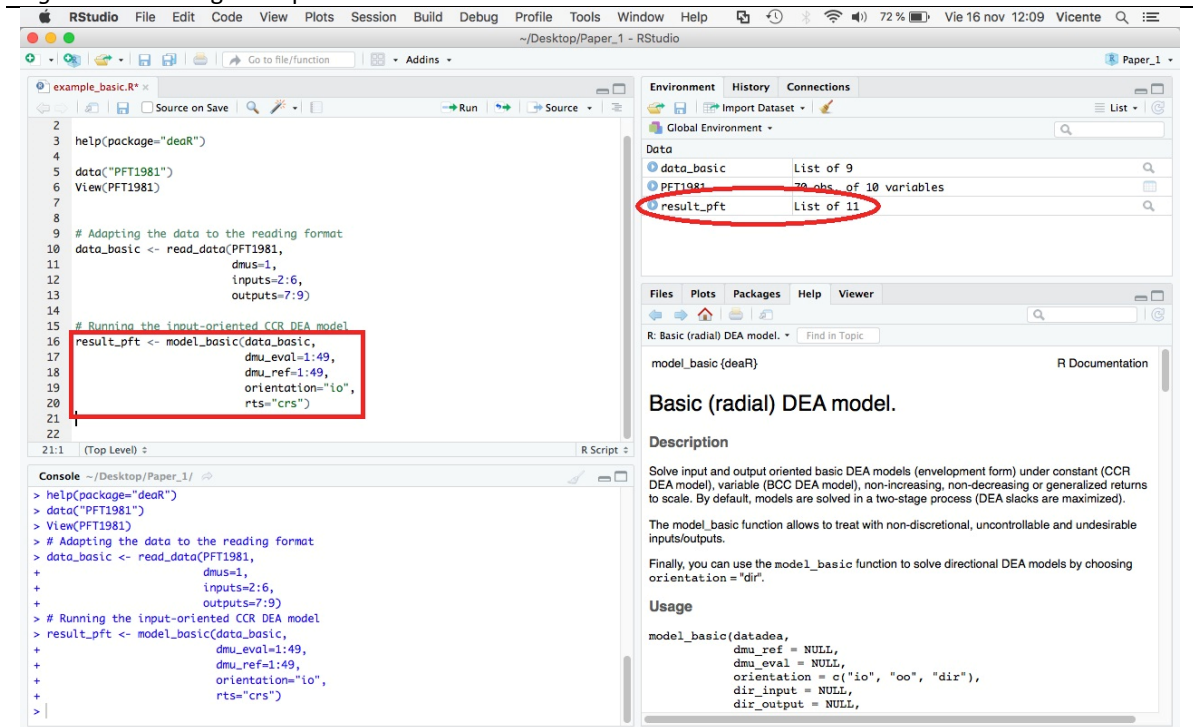


```
result_pft <- model_basci(PFT1981,
                          dmu_ref=1:49,
                          dmu_eval=1:49,
                          orientation="io",
                          rts="crs")
```

and run it (see Figure 37).

Note: We can also select all the instructions and run them instead of doing it instruction by instruction.

Figure 37. Running the input-oriented CCR DEA model for DMUS in PFT.

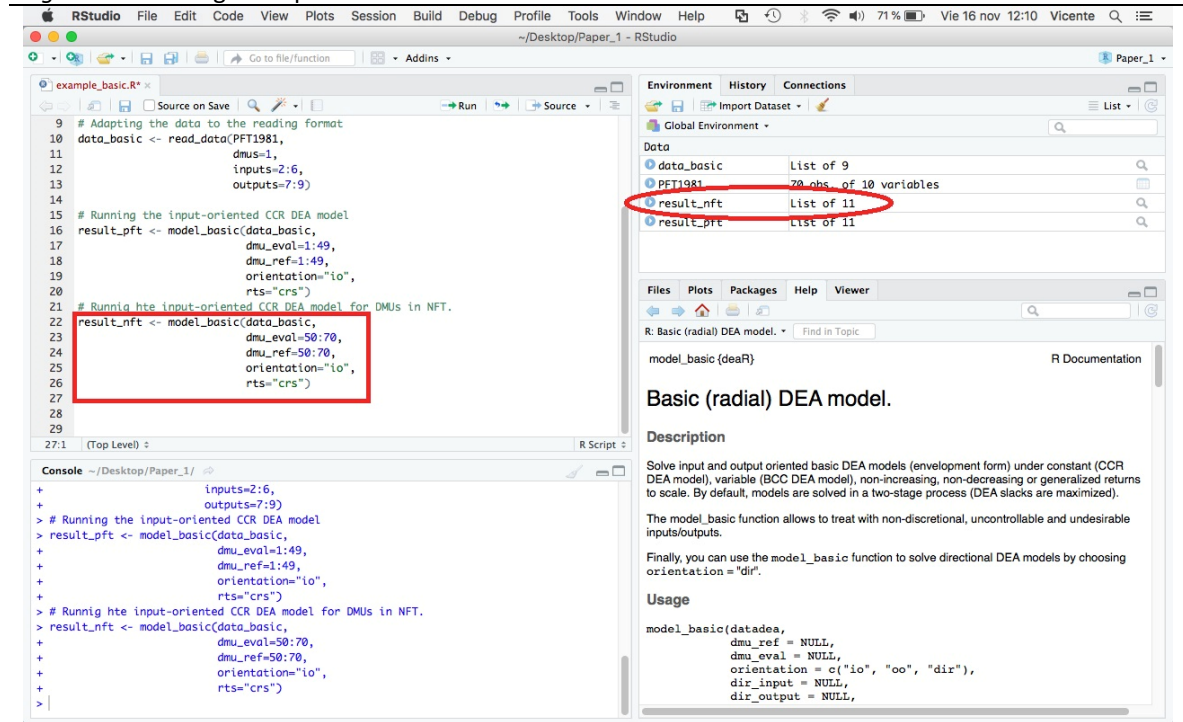


**Example 10.** `model_basci()` function.

At this point, **could you run the CCR input-oriented DEA model for the DMUs in NFT?** (DMUs in NFT are the DMUs from 50 to 70).

*The answer to this question is in next page (Figure 38)*

Figure 38. Running the input-oriented CCR DEA model for DMUs in NFT.

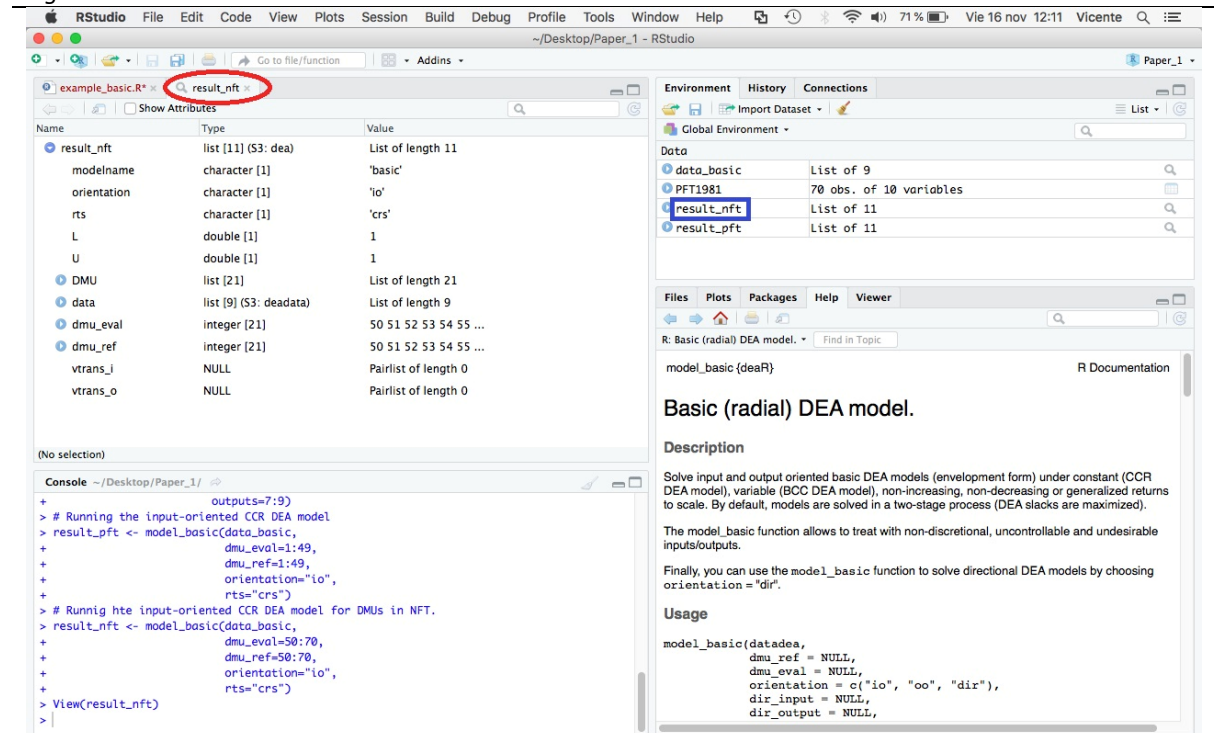


Save the script “*example\_basic*”.

#### 7.4. Extracting the main results.

In Examples 9 and 10, after running the `model_basic()` function, all the results are saved in the objects “*result\_pft*” and “*result\_nft*” respectively. We can see that these objects are a list of 11 components. We can show the content of each list by clicking on the name of the object (see Figure 39) or its structure clicking on the blue arrow (▶).

Figure 39. Structure of the results for a basic DEA model.



To extract the main results of the DEA analysis, **deaR** has a several specific functions. These functions are:

- ✓ **efficiencies()**: To extract the efficiency scores.
- ✓ **slacks()**: To extract input and output slacks.
- ✓ **targets()**: To extract the input and output target values.
- ✓ **lambdas()**: To extract the lambdas (or intensities).
- ✓ **references()**: To extract the reference set for inefficient DMUs.
- ✓ **rts()**: To extract the returns-to-scale.
- ✓ **multipliers()**: To extract the multipliers (or weights) of the multiplier DEA form.

In the following Example 11 we are going to practice with these functions.

### Example 11. Extracting the results.

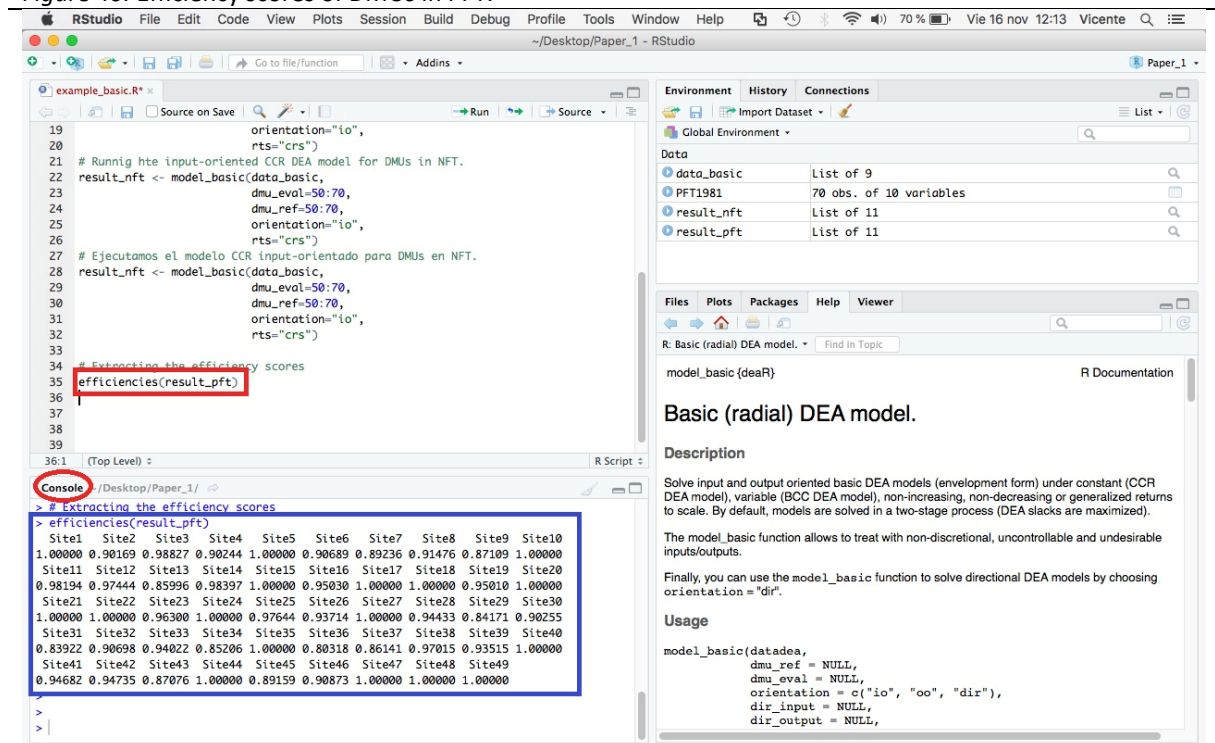
To extract the efficiency scores from “*result\_pft*”, we only have to write into the script the instruction:

**efficiencies(result\_pft)**

and run it.

The efficiency scores will be shown in the *Console*, as the following Figure shows.

Figure 40. Efficiency scores of DMUs in PFT.



Similarly, to extract the reference set of the inefficient DMUs we write and execute the instruction:

**references(result\_pft)**

The remaining functions (`slacks()`, `targets()`, `lambdas()`, `rts()`, and `multipliers()`) work in a similar way. Try and practice with these functions and pay attention to the results obtained.

---

Save the script “*example\_basic..R*”, close the project and quit RStudio.

### 7.5. `summary()` function.

In addition to the previous functions (`efficiencies()`, `lambdas()`, etc.), **deaR** has the `summary()` function, which summarizes the results of the DEA analysis. The `summary()` function serves to summarize the results of both conventional DEA models and fuzzy DEA models. To use this function we have to know that it has the following structure:<sup>19</sup>:

**`summary(objet, exportExcel = TRUE, filename = NULL)`**

The arguments of this function refer to:

- *object*: It is the object to which we have assigned the results of a conventional DEA or fuzzy DEA model.
- *exportExcel*: The default value for this argument is TRUE. Therefore, the `summary()` function will automatically create an Excel file with the main results in the working directory. The user can choose not to create the Excel file (`exportExcel = FALSE`).
- *filename*: By default, the value of the argument is NULL. Therefore, the default name of the Excel file will be: “*ResutsDEAyear-month-dat\_hour:minute:second.xlsx*”. Of course, the user can give other name to the Excel file.

As we have said, the way to use the `summary()` function to summarize the results of a DEA analysis is identical in both conventional DEA models and fuzzy DEA models. The difference is found in the summary of the Excel file generated. That is, the sheets of the Excel workbook created are different depending on the model. Let’s see this in Examples 12 and 13.

---

**Example 12.** Summary of results: conventional DEA model.

Follow the following steps:

1. Open the project “*Paper\_1*”.
2. Create a new script: “*Summary\_DEA*”
3. Load **deaR**.

**Assumption:** Load the dataset provided by **deaR**: “*Hua\_Bian\_2007*”<sup>20</sup>. This dataset has 30 DMUs with 2 inputs (*D-Input1*, *D-Input2*), 2 outputs (*D-Output1*, *D-Output2*) and 1 undesirable output (*UD-Output1*) (in this same order).

We want to get the summary of results from the output-oriented BBC DEA model. As there is an undesirable output, to take this into account in the analysis we will use as the translation parameter the one used by Hua and Bian (2007): `vtrans_o=1500`.

---

<sup>19</sup> For more details: [help\(summary.dea\)](#) o [help\(summary.dea\\_fuzzy\)](#).

<sup>20</sup> Hua Z.; Bian Y. (2007). DEA with Undesirable Factors. In: Zhu J., Cook W.D. (eds) Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer, Boston, MA. [https://doi.org/10.1007/978-0-387-71607-7\\_6](https://doi.org/10.1007/978-0-387-71607-7_6)

**Important:** Remember the steps to use **deaR**.

Step 1. Load the data.

Step 2. Adapt the data to the format used by **deaR**.

Step 3. Run the DEA model.

Step 4. Extract the results.

**Try it!!!**

(the solution is shown in the next page)



**Solution:** As we can see in Figure 41, to solve Example 12 we have to write into the script “Summary\_DEA” the following instructions:

Step 1. Load the data:

```
data("Hua_Bian_2007")
```

Step 2. Adapt the data:

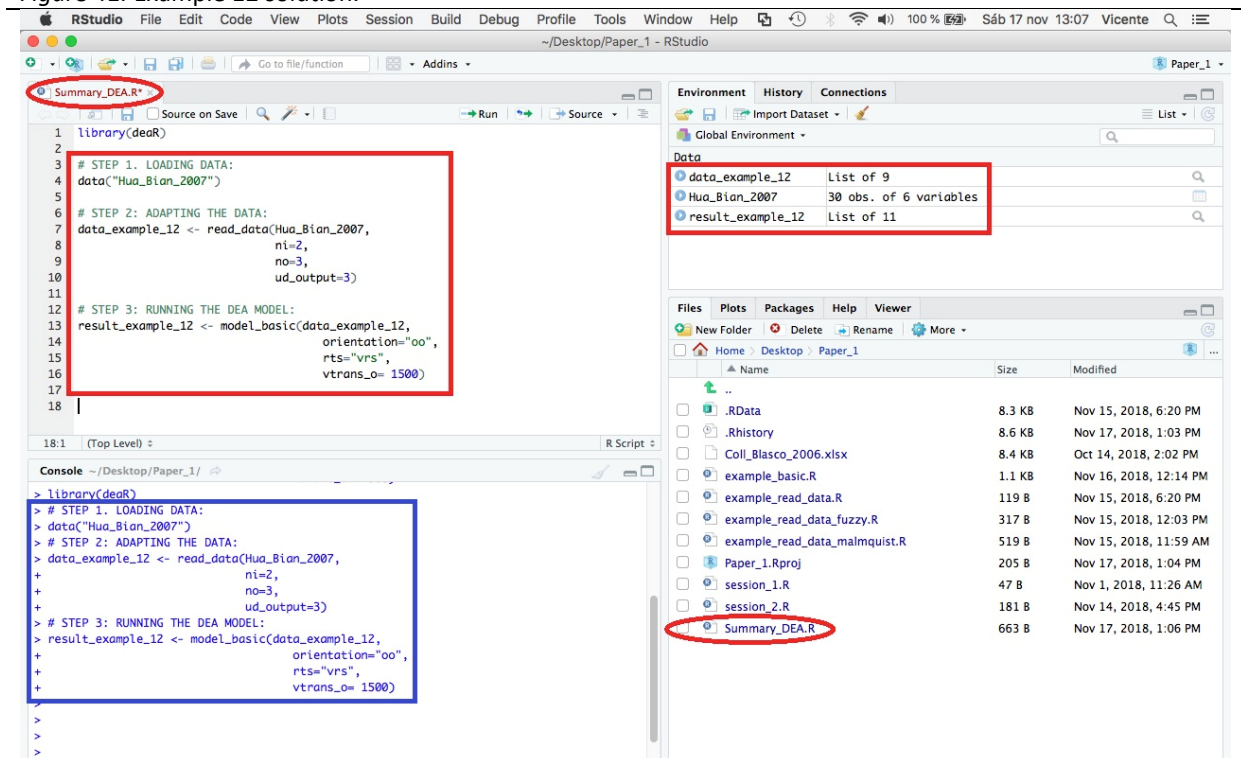
```
data_example_12 <- read_data(Hua_Bian_2007,
                             ni=2,
                             no=3,
                             ud_output=3)
```

**Note:** We have 3 outputs (no=3) and the third output is the undesirable output (ud\_output=3)

Step 3. Run the DEA model:

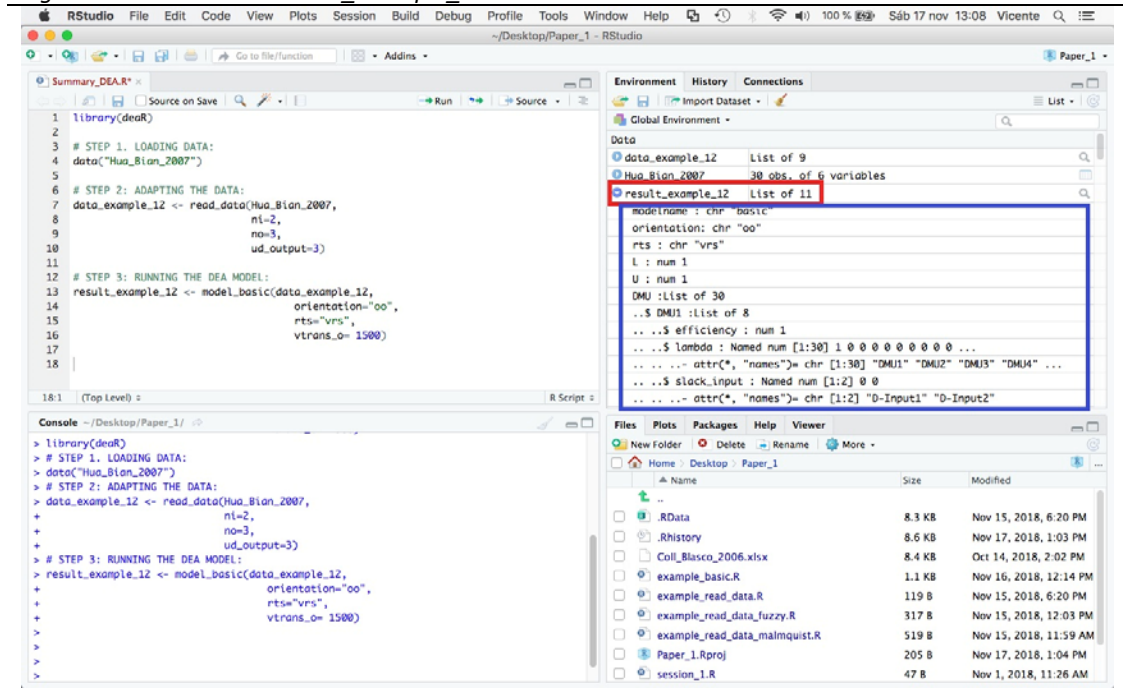
```
result_example_12 <- model_basic(data_example_12,
                                orientation="oo",
                                rts="vrs",
                                vtrans_o= 1500)
```

Figure 41. Example 12 solution.



All the results of the output-oriented BCC DEA model that we have executed can be found in the object "result\_example\_12", which is a list of 11 components (see Figure 42).

Figure 42. Structure of "result\_example\_12".



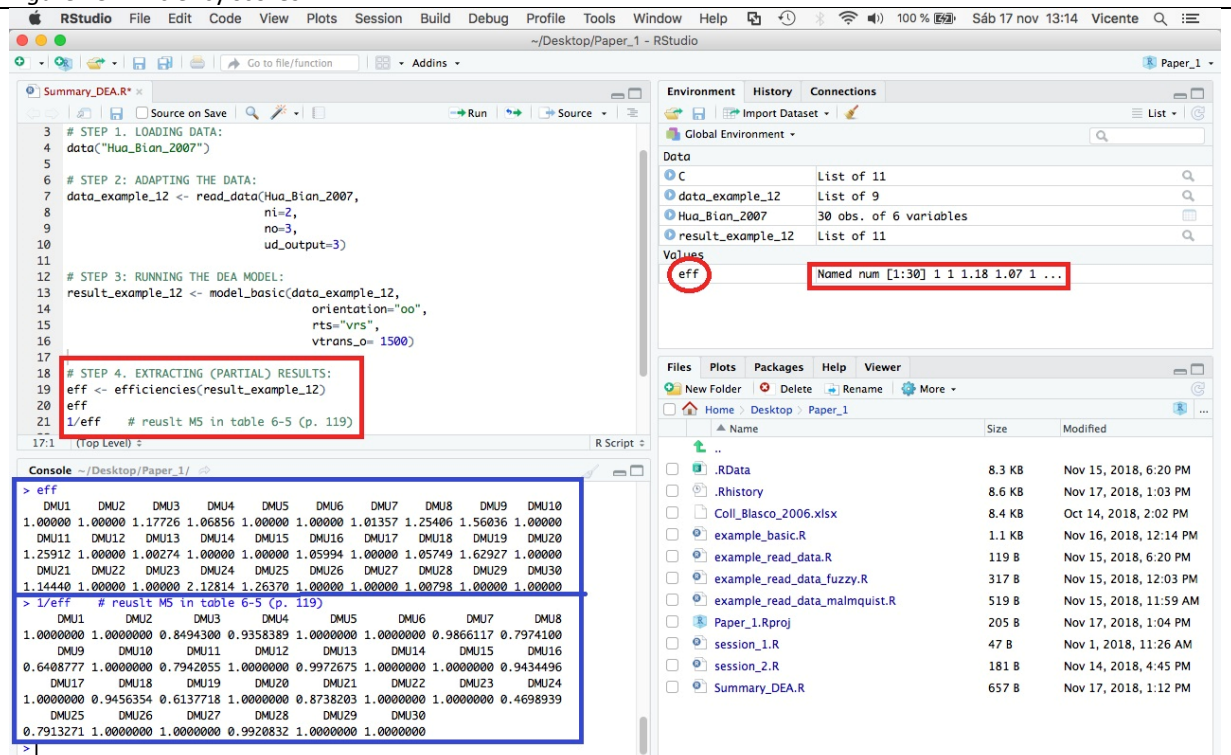
At this point, we can extract the partial results with the functions: `efficiencies()`, `lambdas()`, `multipliers()`, `rts()`, `references()`, `slacks()`, `targets()`; and the summary of results with the `summary()` function.

As the Figure 43 shows, we extract the efficiency scores of the DMUs with the `efficiencies()` function. The efficiencies are listed in the *Console* and assigned to the object "eff". To obtain the same results as those obtained by Hua and Bian (2007), we write into the script:

**1/eff**

and run the instruction.

Figure 43. Efficiency scores.

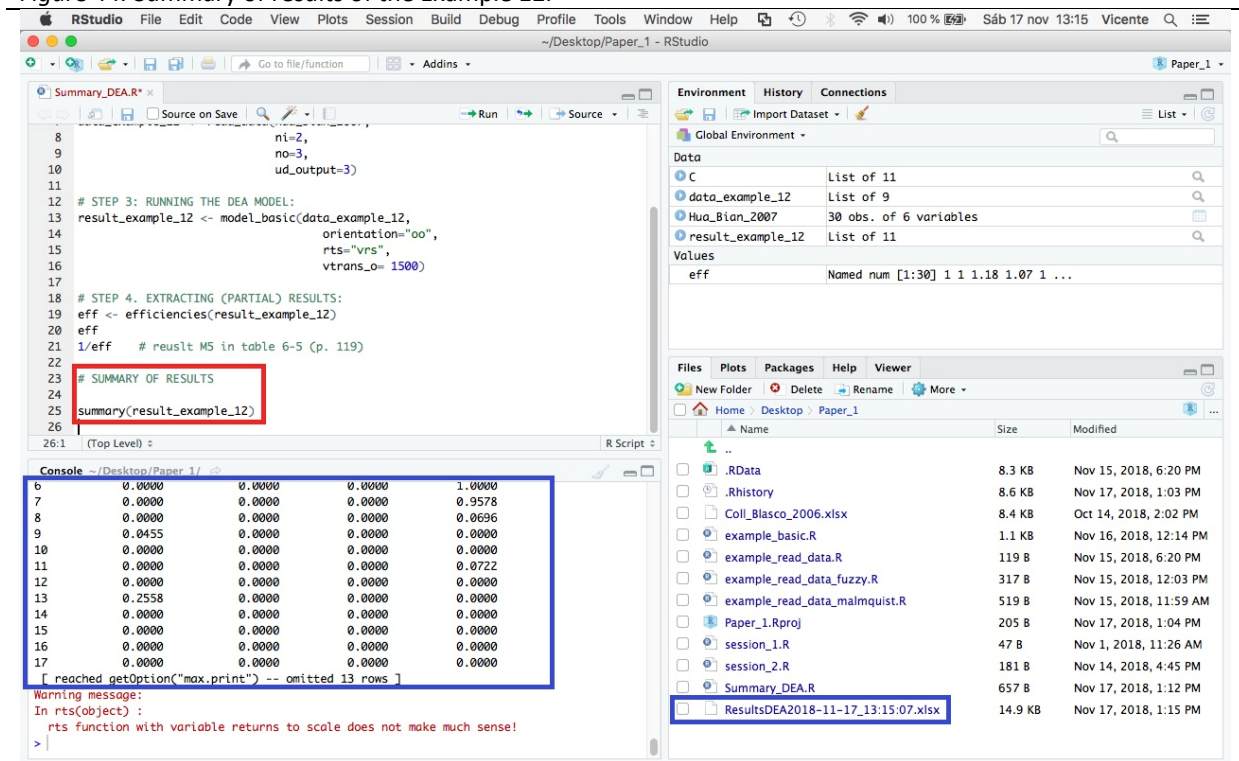


To get a summary of all the results obtained by running the output-oriented BCC DEA model, we use the `summary()` function. We write into the script "Summary\_DEA":

```
summary(result_example_12)
```

All results are dumped on the *Console*. Although there are not many DMUs (only 30), there are many results. It is not practical to use the `summary()` function to see the results on the screen. However, as we can see in Figure 44, **dearR** has also created an Excel file that has all these results. Notice the name that has been given to the file by default.

Figure 44. Summary of results of the Example 12.



The screenshot shows the RStudio interface with a script editor, environment pane, and console. The script editor contains the following code:

```

8         ni=2,
9         no=3,
10        ud_output=3)
11
12 # STEP 3: RUNNING THE DEA MODEL:
13 result_example_12 <- model_basic(data_example_12,
14                                orientation="oo",
15                                rts="vrs",
16                                vtrans_o= 1500)
17
18 # STEP 4. EXTRACTING (PARTIAL) RESULTS:
19 eff <- efficiencies(result_example_12)
20 eff
21 1/eff # result MS in table 6-5 (p. 119)
22
23 # SUMMARY OF RESULTS
24 summary(result_example_12)
25
26

```

The console output shows the following table:

6	0.0000	0.0000	0.0000	1.0000
7	0.0000	0.0000	0.0000	0.9578
8	0.0000	0.0000	0.0000	0.0696
9	0.0455	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000	0.0000
11	0.0000	0.0000	0.0000	0.0722
12	0.0000	0.0000	0.0000	0.0000
13	0.2558	0.0000	0.0000	0.0000
14	0.0000	0.0000	0.0000	0.0000
15	0.0000	0.0000	0.0000	0.0000
16	0.0000	0.0000	0.0000	0.0000
17	0.0000	0.0000	0.0000	0.0000

The console also shows a warning message:

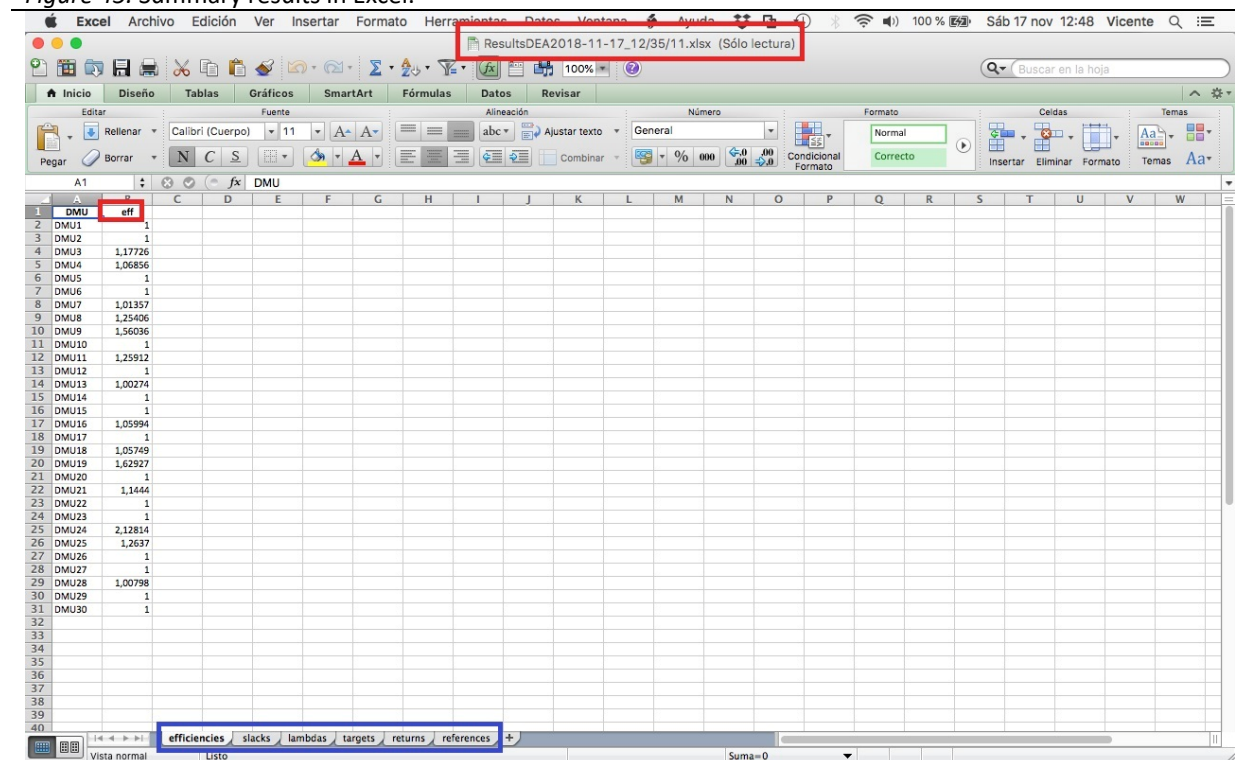
```
Warning message:
In rts(object) :
  rts function with variable returns to scale does not make much sense!
```

The file explorer shows the following files:

Name	Size	Modified
..		
.RData	8.3 KB	Nov 15, 2018, 6:20 PM
.Rhistory	8.6 KB	Nov 17, 2018, 1:03 PM
Coll_Blasco_2006.xlsx	8.4 KB	Oct 14, 2018, 2:02 PM
example_basic.R	1.1 KB	Nov 16, 2018, 12:14 PM
example_read_data.R	119 B	Nov 15, 2018, 6:20 PM
example_read_data_fuzzy.R	317 B	Nov 15, 2018, 12:03 PM
example_read_data_malmquist.R	519 B	Nov 15, 2018, 11:59 AM
Paper_1.Rproj	205 B	Nov 17, 2018, 1:04 PM
session_1.R	47 B	Nov 1, 2018, 11:26 AM
session_2.R	181 B	Nov 14, 2018, 4:45 PM
Summary_DEA.R	657 B	Nov 17, 2018, 1:12 PM
ResultsDEA2018-11-17_13:15:07.xlsx	14.9 KB	Nov 17, 2018, 1:15 PM

Open the Excel file to see how the results are shown (see Figure 45). To do this, we click on the file "ResultsDEAyear-month-dat\_hour:minute:second.xlsx" and select the option "View file".

Figure 45. Summary results in Excel.



DMU	eff
DMU1	1
DMU2	1
DMU3	1,17726
DMU4	1,06856
DMU5	1
DMU6	1
DMU7	1,01357
DMU8	1,25406
DMU9	1,56036
DMU10	1
DMU11	1,25912
DMU12	1
DMU13	1,00274
DMU14	1
DMU15	1
DMU16	1,05994
DMU17	1
DMU18	1,05749
DMU19	1,62927
DMU20	1
DMU21	1,1444
DMU22	1
DMU23	1
DMU24	2,12814
DMU25	1,2637
DMU26	1
DMU27	1
DMU28	1,00798
DMU29	1
DMU30	1

We save the script "summary\_DEA".

Now, in Example 13 we are going to replicate the results in León, et al. (2003)<sup>21</sup>.

**Example 13.** Summary of results: Possibilistic fuzzy DEA model.

León et al. (2013) use possibilistic programming techniques to measure the efficiency based on the envelopment form of the input-oriented BCC model. Data used by the authors in their article can be found in the dataset "Leon2003", which is provided by deaR.

We create a new script: "Summary\_DEA\_fuzzy". To replicate the results that the authors show in Table 2 (page 419), we write into the script:

```
data("Leon2003")
data_example <- read_data_fuzzy(Leon2003,
                                dmus = 1,
                                inputs.mL = 2,
                                inputs.dL = 3,
                                outputs.mL = 4,
                                outputs.dL = 5)

result <- modelfuzzy_possibilistic(data_example,
                                   h = seq(0, 1, by = 0.1),
                                   orientation = "io",
                                   rts = "vrs")

efficiencies(result)
```

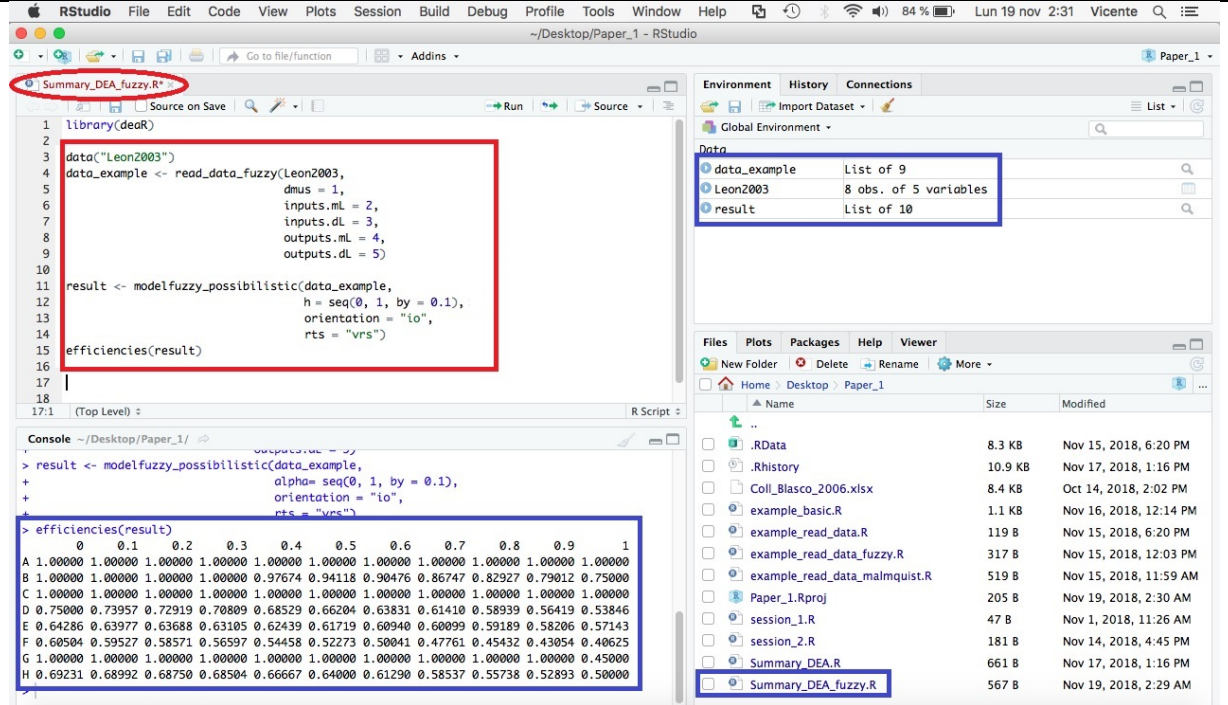
<sup>21</sup> León, T.; Liern, V. Ruiz, J.; Sirvent, I. (2003). "A Possibilistic Programming Approach to the Assessment of Efficiency with DEA Models", Fuzzy Sets and Systems, 139, 407–419. [https://doi.org/10.1016/S0165-0114\(02\)00608-5](https://doi.org/10.1016/S0165-0114(02)00608-5)



**Note:**  $h = \text{seq}(0, 1, \text{by} = 0.1)$  generates a sequence of values for the different levels of possibility:  $h = (0, 0.1, 0.2, \dots, 1)$ .

We run the instructions. The efficiency scores of the BCC input-oriented model will be shown for the different levels of possibility  $h$  in the *Console* (see Figure 46)

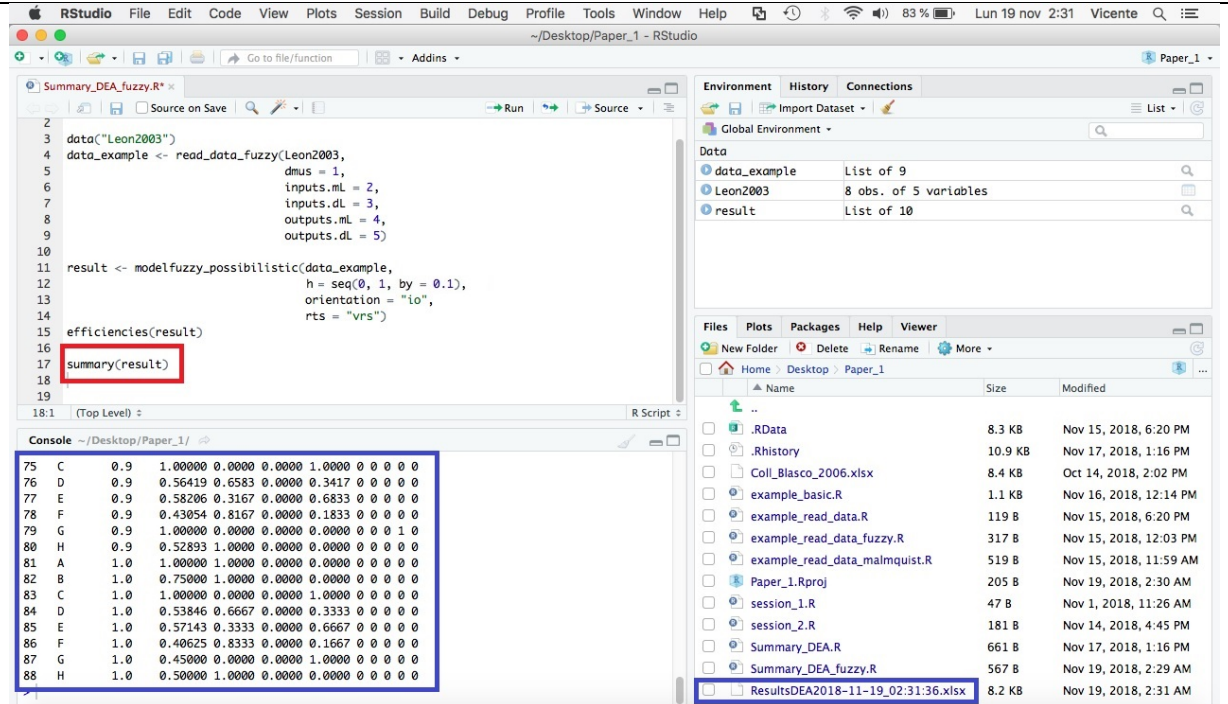
Figure 46. Efficiency scores for different  $h$ -levels.



To get a summary of the results on the screen (see Figure 47) and export them to an Excel file, we write the following instruction into the script and run it:

`summary(result)`

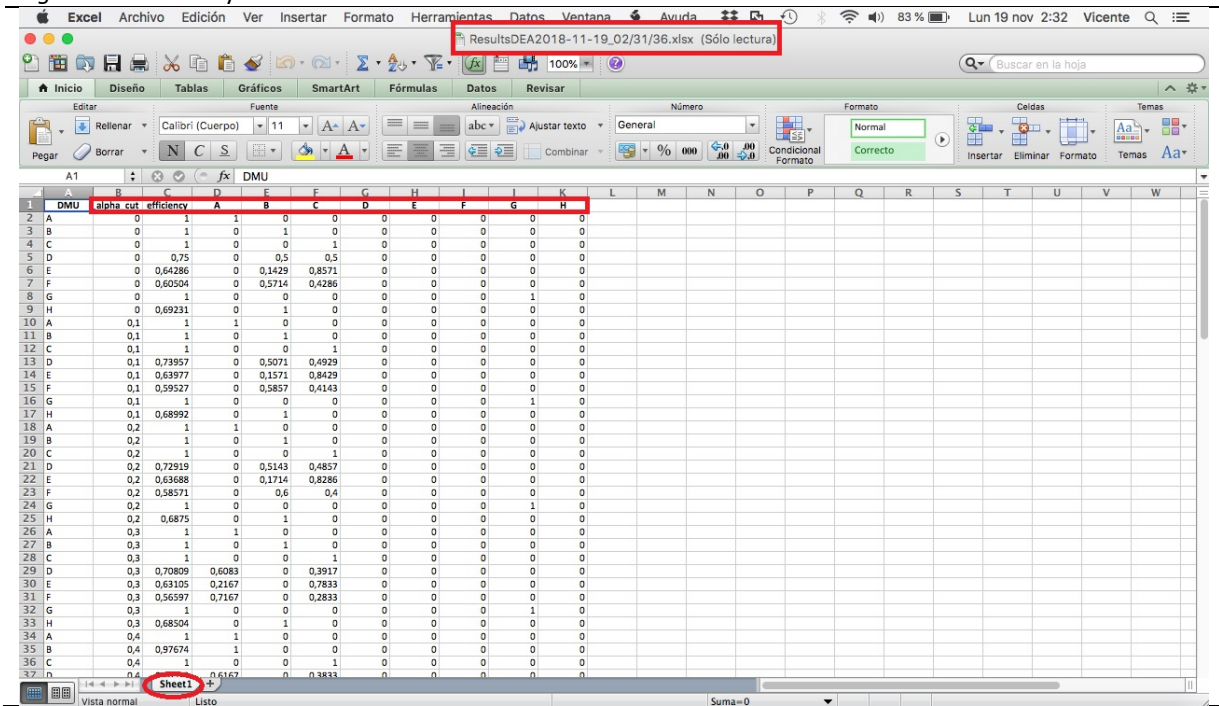
Figure 47. Summary of results for the possibilistic input-oriented BCC DEA model.





For the possibilistic model, the summary of results consists of: (1) the efficiency scores and (2) the reference sets. These are the results that are exported to the Excel workbook (see Figure 48).

Figure 48. Summary results in Excel.



DMU	alpha cut	efficiency	DMU	DMU	DMU	DMU	DMU	DMU	DMU
2	A	0	1	0	0	0	0	0	0
3	B	0	1	0	1	0	0	0	0
4	C	0	1	0	0	1	0	0	0
5	D	0	0.75	0	0.5	0.5	0	0	0
6	E	0	0.64286	0	0.1429	0.8571	0	0	0
7	F	0	0.60504	0	0.5714	0.4286	0	0	0
8	G	0	1	0	0	0	0	0	1
9	H	0	0.69231	0	1	0	0	0	0
10	A	0.1	1	1	0	0	0	0	0
11	B	0.1	1	0	1	0	0	0	0
12	C	0.1	1	0	0	1	0	0	0
13	D	0.1	0.73857	0	0.5071	0.4929	0	0	0
14	E	0.1	0.63977	0	0.1571	0.8429	0	0	0
15	F	0.1	0.59527	0	0.5857	0.4143	0	0	0
16	G	0.1	1	0	0	0	0	0	1
17	H	0.1	0.68992	0	1	0	0	0	0
18	A	0.2	1	1	0	0	0	0	0
19	B	0.2	1	0	1	0	0	0	0
20	C	0.2	1	0	0	1	0	0	0
21	D	0.2	0.72919	0	0.5143	0.4857	0	0	0
22	E	0.2	0.63688	0	0.1714	0.8286	0	0	0
23	F	0.2	0.58571	0	0.6	0.4	0	0	0
24	G	0.2	1	0	0	0	0	0	1
25	H	0.2	0.6875	0	1	0	0	0	0
26	A	0.3	1	1	0	0	0	0	0
27	B	0.3	1	0	1	0	0	0	0
28	C	0.3	1	0	0	1	0	0	0
29	D	0.3	0.70809	0.6083	0	0.3917	0	0	0
30	E	0.3	0.63105	0.2167	0	0.7833	0	0	0
31	F	0.3	0.56597	0.7167	0	0.2833	0	0	0
32	G	0.3	1	0	0	0	0	0	1
33	H	0.3	0.68504	0	1	0	0	0	0
34	A	0.4	1	1	0	0	0	0	0
35	B	0.4	0.97674	1	0	0	0	0	0
36	C	0.4	1	0	0	1	0	0	0
37	D	0.4	0.6167	0	0.3833	0	0	0	0

We save the script “*Summary\_DEA\_fuzzy*”.

## 7.6. Graphical results. The `plot()` function.

With the `plot()` function we can obtain some plots of the results from a conventional DEA model, Malmquist index or cross-efficiency model. This is shown in Examples 14, 15, and 16 respectively.

The `plot()` function is used in the following way:

`plot(x)`

where `x` is the object in which the result of a DEA model has been stored.

After using the `plot()` function, we will get the plot in the *Viewer* (right bottom window). We can save a plot by clicking on the *Export* tag of the *Viewer*. It is possible to save the plot as an image in format: “*png*”, “*jpeg*” or “*tiff*”. We can copy the plot to the system clipboard and then paste it, for example, in a word document.

Graphs for fuzzy DEA models are not available still in this version of **deaR**.

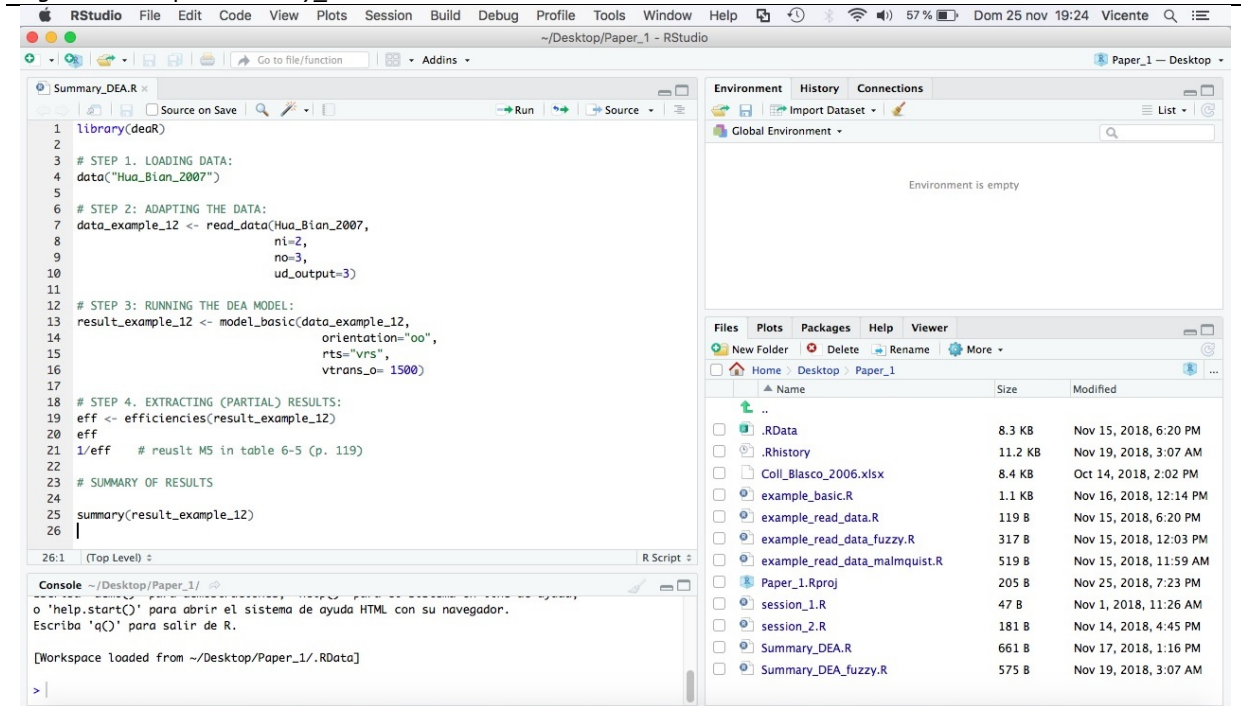
We are currently working to improve the graphical results in **deaR**, which will be included in the next version.

### Example 14. Plot: Basic DEA model.

We open the script “*Summary\_DEA*”. If we have closed the working session, we need to open the project “*Paper\_1*”, load **deaR**, and open the script.

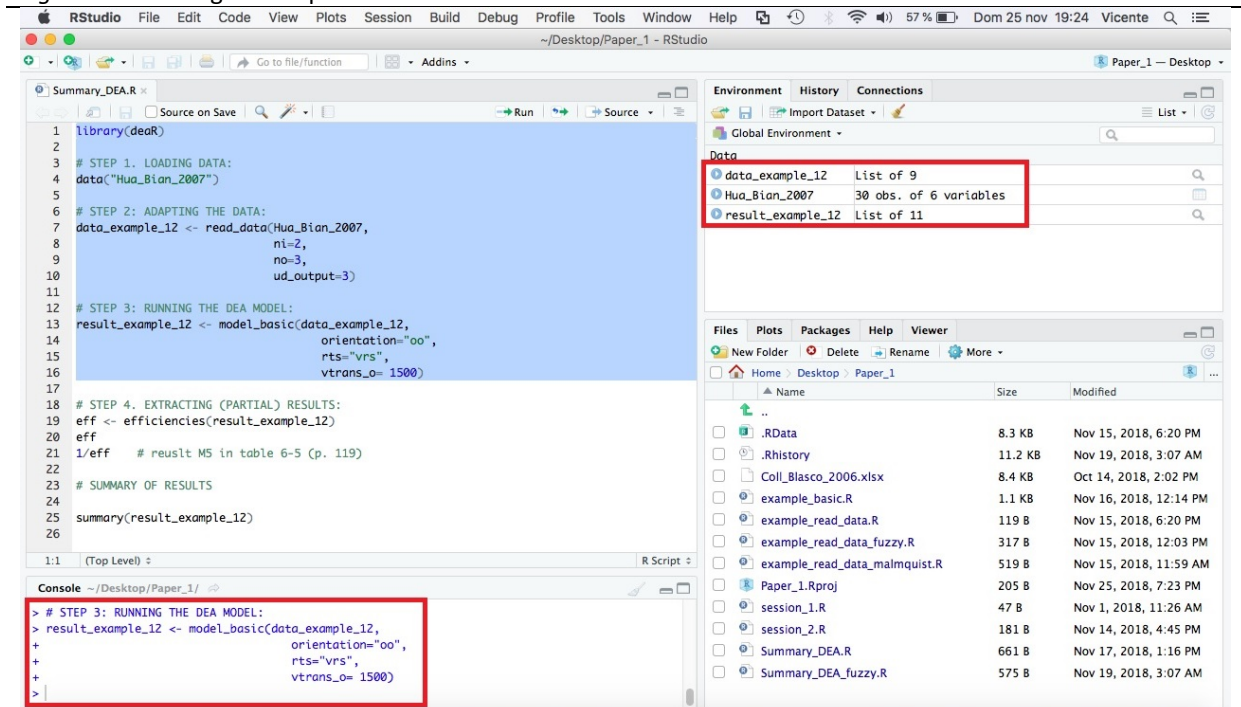
Our working session should be similar to the one shown in Figure 49.

Figure 49. Script “Summary\_DEA”.



As Figure 50 shows, we select from line 1, **library(dear)**, to line 16 and run the selection. Immediately, three objects are listed in the *Environment*: “*data\_example\_12*”, “*Hua\_Bian\_2007*” and “*result\_example\_12*”.

Figure 50. Running the script.



The results of the DEA model are assigned to (or stored in) the object “*result\_example\_12*”. To plot some of these results, we will write the following instruction on the line 27 of the script:

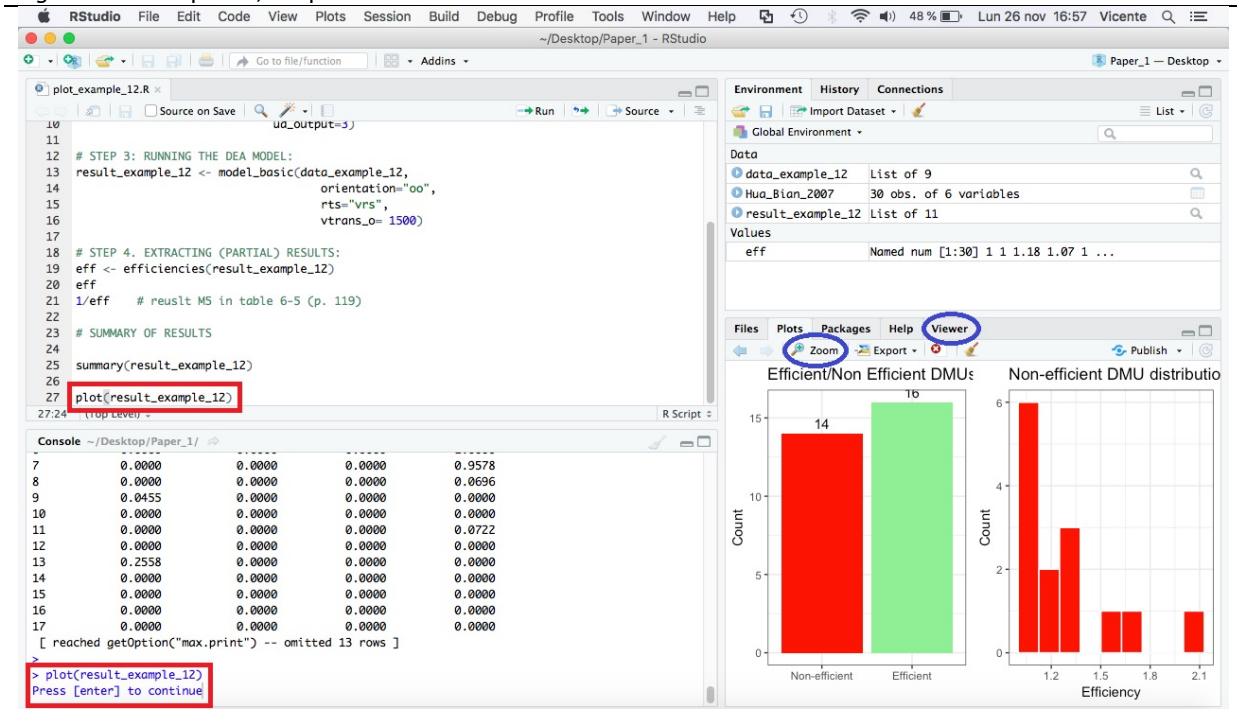
**plot(result\_example\_12)**

and run it. The following message will be displayed in the *Console*:

**Press [enter] to continue**

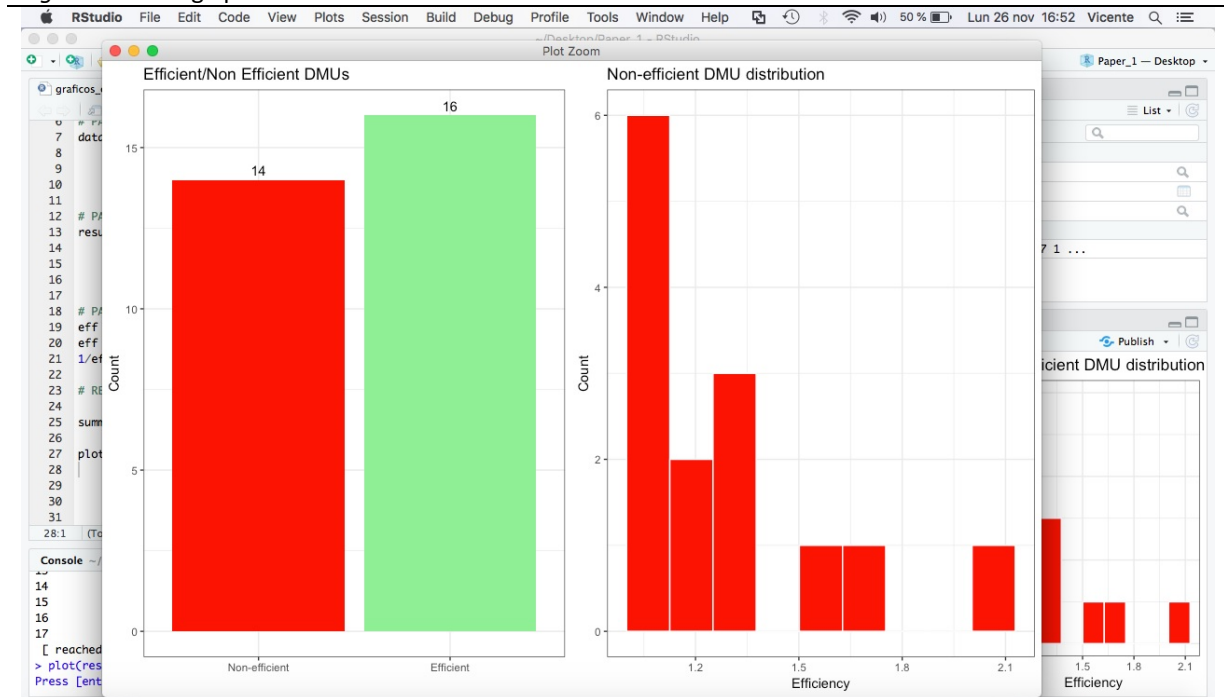
and a plot will be shown in the *Viewer* tag (right bottom window), as we can see in Figure 51.

Figure 51. Example 12, Graphic 1.



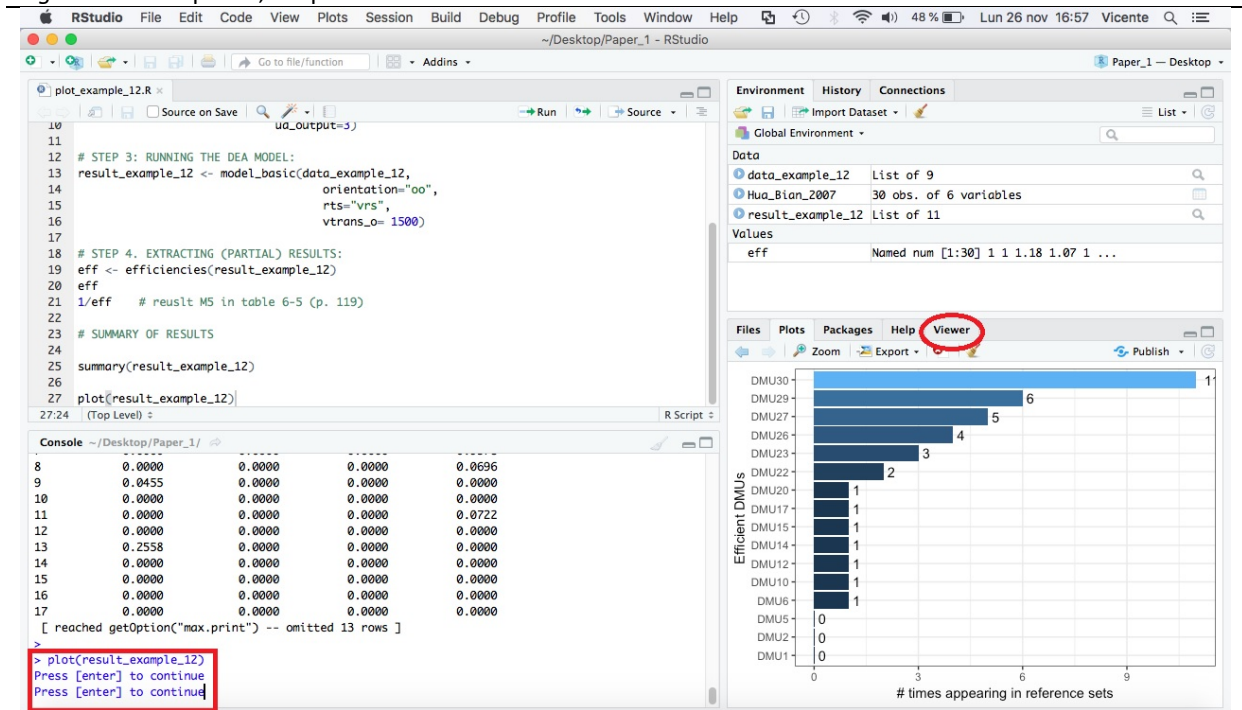
This Graphic 1 shows the number of efficient and inefficient DMUs (right graph), and the distribution of the efficiency score of the latter as well (left graph). If we click on the Zoom bottom, the plot will be enlarged (see Figure 52).

Figura 52. Enlarge plot.



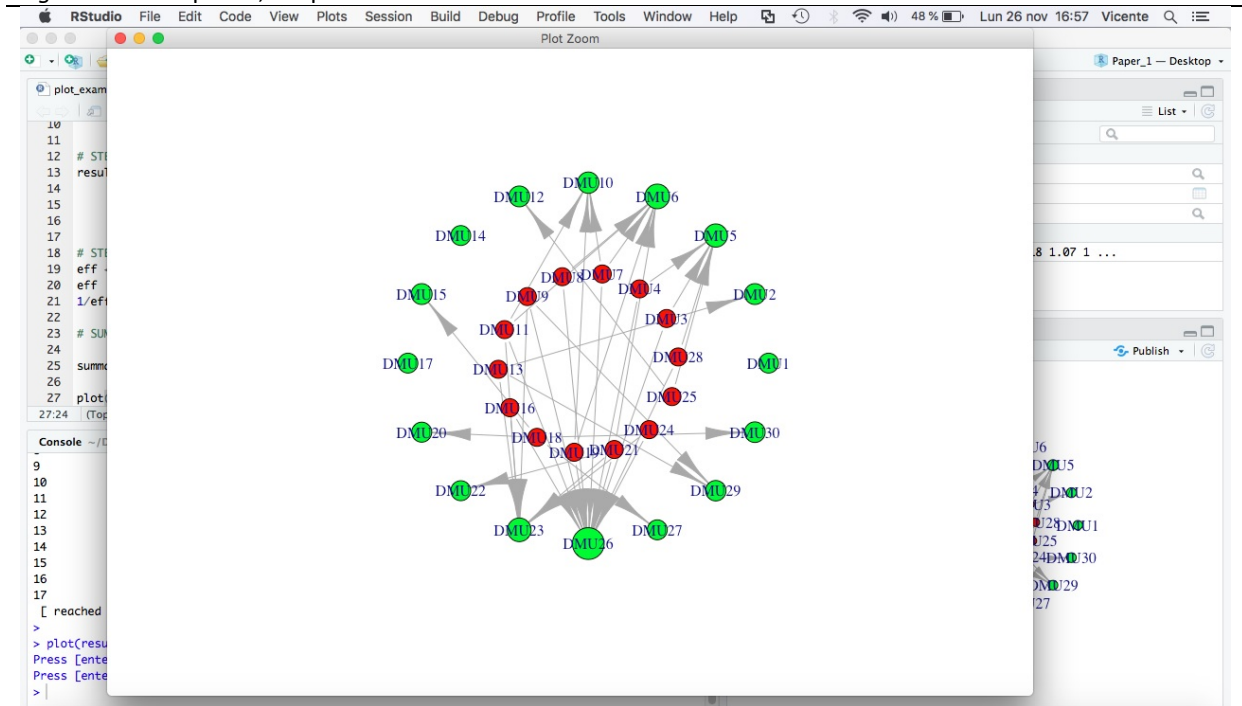
Now, if we press the *Enter* key, a second graphic will be displayed. On this occasion, what is depicted is the number of times an efficient DMU is part of the inefficient DMUs reference set.

Figure 53. Example 12, Graphic 2.



Finally, by pressing the *Enter* key again we can see the last graphic (see Figure 54).

Figure 54. Example 12, Graphic 3.



In Figure 54 we see a network graph in which the green circles represent the efficient DMUs and the red circles the inefficient ones. This graph shows how inefficient DMUs are related to efficient DMUs, trying to convey the idea that the efficient ones form the efficient frontier. In addition, notice that not all green circles have the same size. In this case, the size



of the circle aims to convey the idea of how important is the efficient DMU for the set of inefficient DMUs.

We save the script as: “*plot\_example\_12*”.

**Example 15.** Plot: Malmquist index.

We create a new script with the name “*plot\_malmquist*”. If we have closed the working session, we need to open the project “*Paper\_1*”, load **dear**, and create the script.

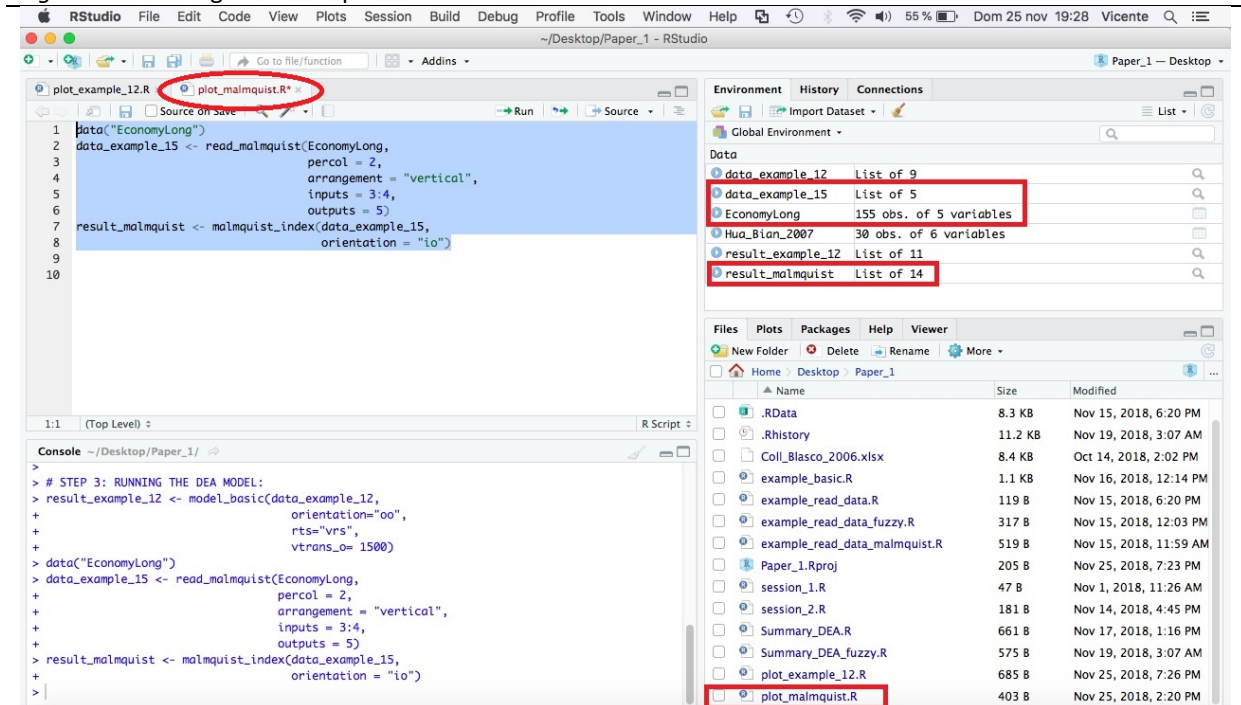
Now, we are going to run the Malmquist index. For that, we write into the script:

```
data("EconomyLong")
data_example_15 <- read_malmquist(EconomyLong,
                                percol = 2,
                                arrangement = "vertical",
                                inputs = 3:4,
                                outputs = 5)
result_malmquist <- malmquist_index(data_example_15,
                                    orientation = "io")
```

**Note:** The data are in long format.

We run the instructions of the script (see Figure 55).

Figure 55. Running the Malmquist index.



The results of the Malmquist index are stored in the object “*result\_malmquist*”. Therefore, to get the plots we write into the script:

```
plot(result_malmquist)
```

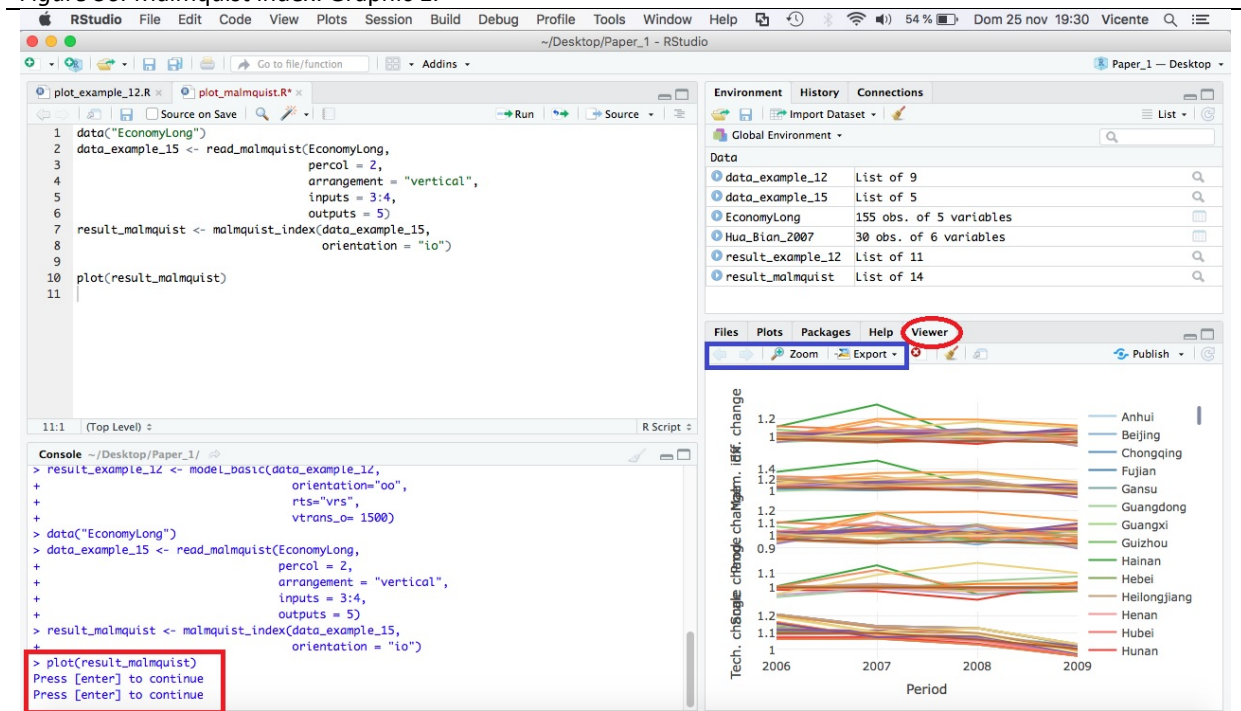


and run the instruction. The following message will be displayed in the *Console*:

**Press [enter] to continue**

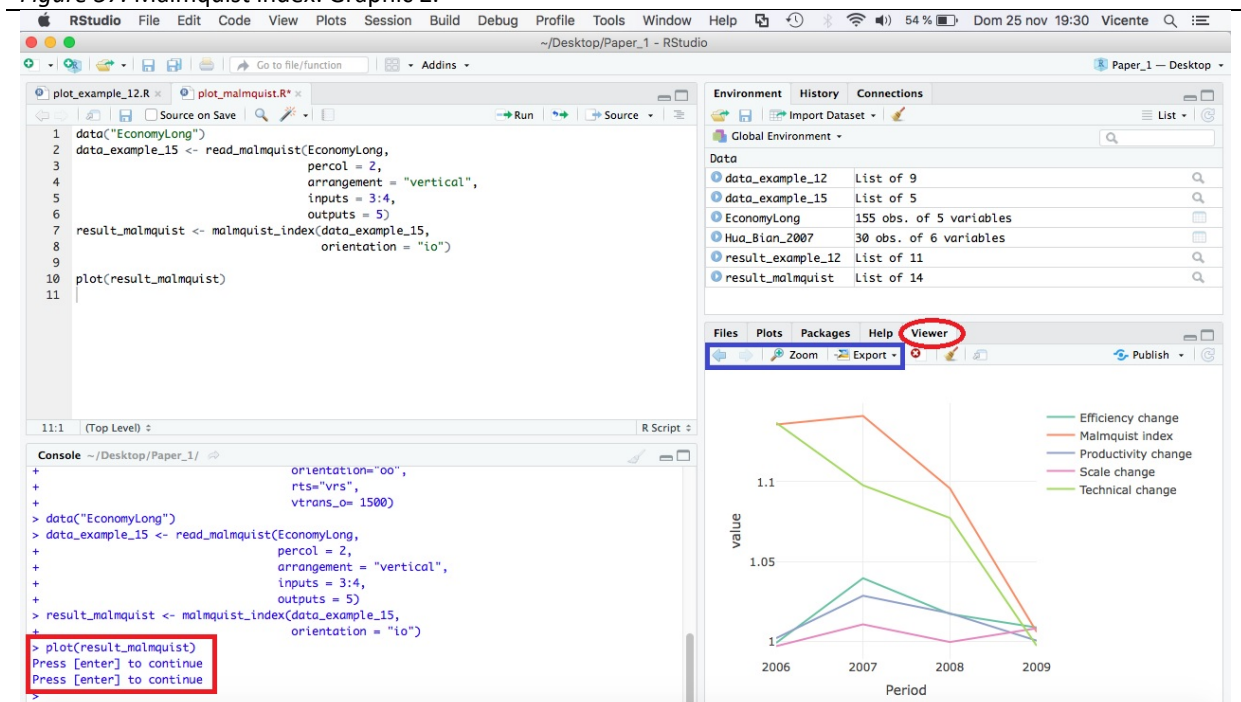
By pressing the *Enter* key a plot will be shown in the *Viewer* (right bottom window) (see Figure 56). We can click on the *Zoom* to see the plot better. In this first plot, for each DMU the Malmquist index and its components are drawn over the years. If there are many DMUs you will not probably see them very well. However, it is possible to select the DMUS the DMUs of your interest to focus on them (and the other DMUs in the graph will disappear).

Figure 56. Malmquist index: Graphic 1.



If we press the *Enter* key (in the *Console*) again, a second plot is created (see Figure 57).

Figure 57. Malmquist index: Graphic 2.



On this occasion, the geometric means of Malmquist index and its components are drawn over the years. We can enlarge the plot by clicking on the *Zoom* and go back (or go forward) by clicking on the arrows. Here, it is possible to select the Malmquist index components to be shown.

We save “*plot\_malmquist*”.

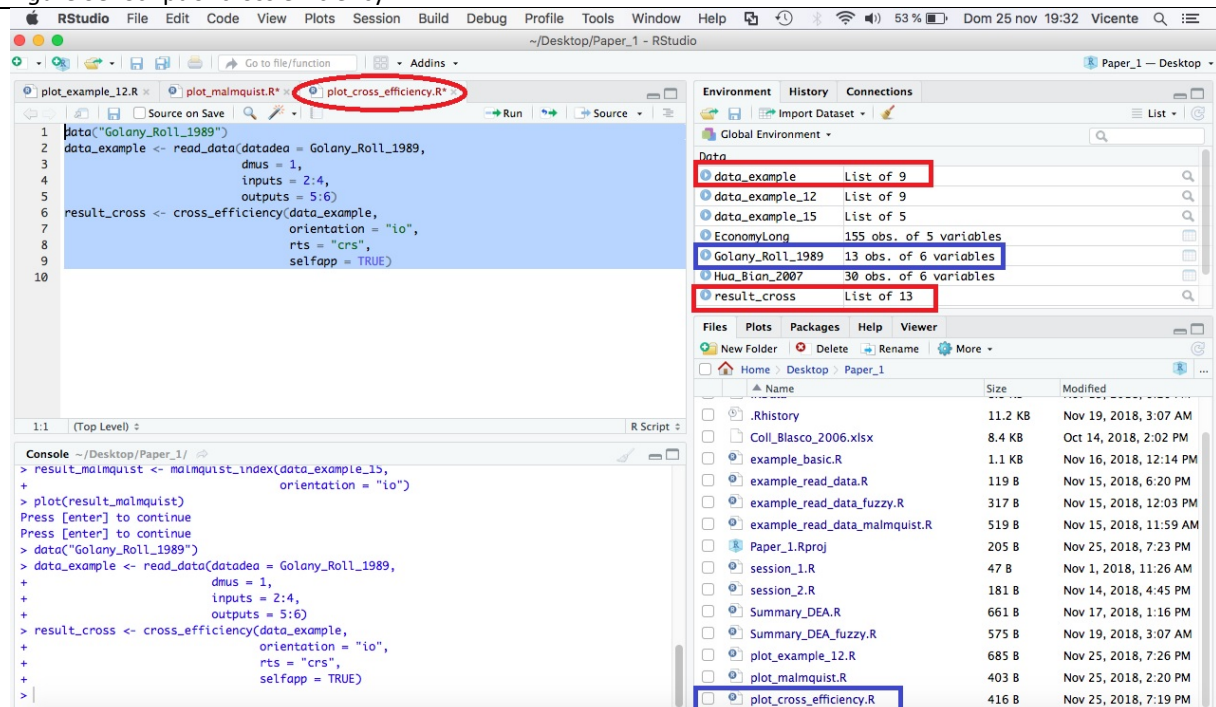
**Example 16.** Plot: Cross efficiency.

We create the new script “*plot\_cross\_efficiency*”. If we have closed the working session, we need to open the project “*Paper\_1*”, load **deaR**, and create the script.

We write and run the followings instructions (see Figure 58):

```
data("Golany_Roll_1989")
data_example <- read_data(datadea = Golany_Roll_1989,
                          dmus = 1,
                          inputs = 2:4,
                          outputs = 5:6)
result_cross <- cross_efficiency(data_example,
                                orientation = "io",
                                rts = "crs",
                                selfapp = TRUE)
```

Figure 58. Script of cross efficiency.

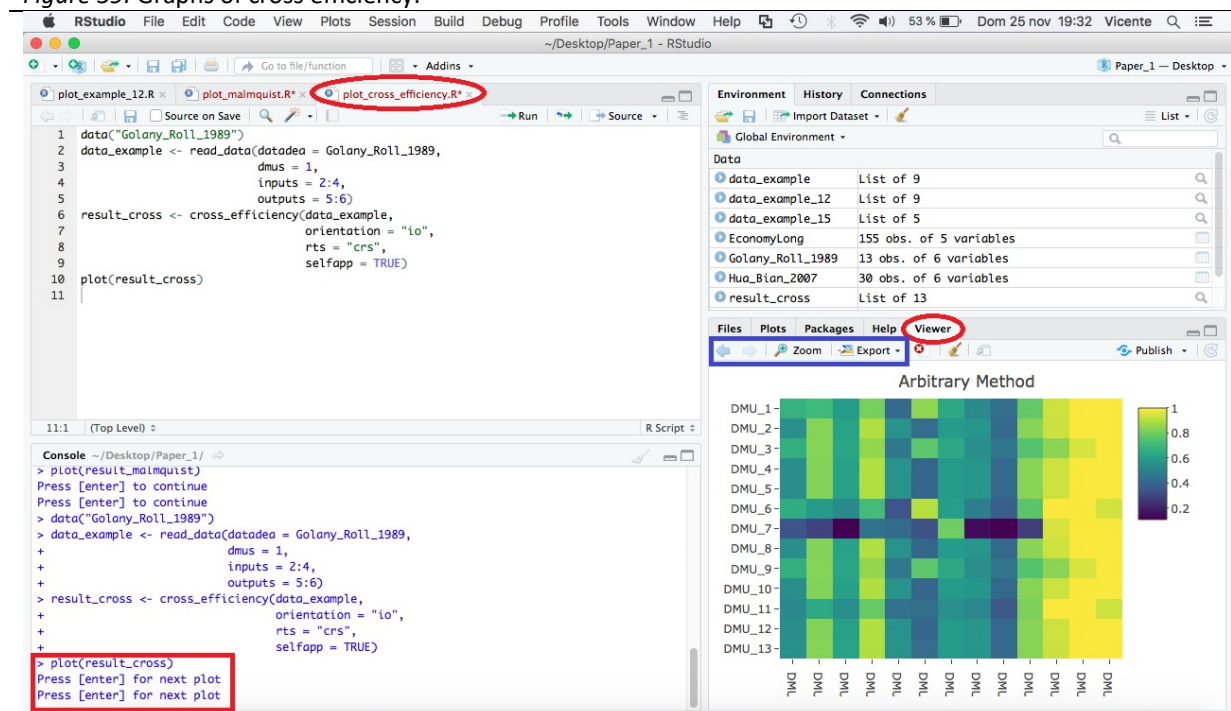


In the object “*result\_cross*” are stored the results of the cross efficiency corresponding to the models: arbitrary, benevolent and aggressive. All these results are drawn like a heat map by using the `plot()` function. For that, we write into the script:

```
plot(result_cross)
```

and press *Enter* in the *Console* to display the different graphs. The result should be similar to the one shown in Figure 59.

Figure 59. Graphs of cross efficiency.



We save the script "*plot\_cross\_efficiency*", close the project and quit RStudio.

## Acknowledgements

The authors wish to thank Dra. Cristina Pardo-Garcia for her inestimable assistance reviewing this tutorial.

We hope you find **deaR** useful and you use it in your future research and teaching.

Any comments and suggestions to improve **deaR** will be really appreciated.