Getting started with R and R studio

Elaborated: Viet Nguyen
Master student in Forest Information Technology
Faculty of Forest and Environment
Eberswalde University for Sustainable Development

Goal: This tutorial aims to give a short introduction to the R programming language and R studio application as an integrated development environment for R. A step-by-step guide to setup R and R studio was also given, the environment that will be further used for LiDAR point cloud processing.

Contents

1. What is R?

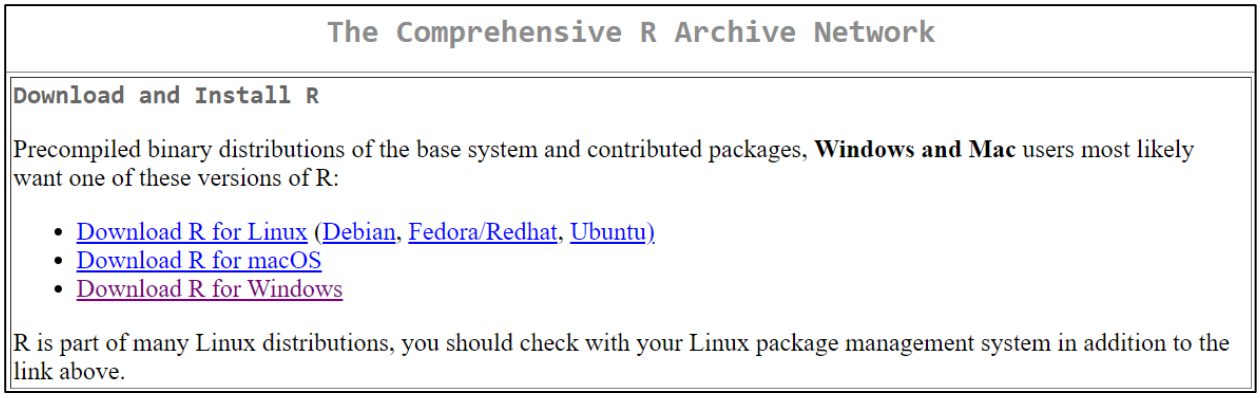
R is not a programming language like C or Java. It was not created by software engineers for software development. Instead, it was developed by statisticians as an interactive environment for data analysis. However, like in other programming languages, you can save your work as scripts that can be easily executed at any moment. These scripts serve as a record of the analysis you performed, a key feature that facilitates reproducible work (Irizarry, 2021).

Some attractive features of R:

- R is free and open source
- R is available for Windows, Macintosh, UNIX/Linux
- R has an extensive help system and a lot of documentation available
- R has huge graphical possibilities
- R has a lot of built-in statistical functions
- There is a large, active support community of R users
- Scripts and data objects can be shared seamlessly across platforms.

2. R downloading

R is available to download from the official website.


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 (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- 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.

The version of R to download depends on your operating system (R 4.0.2 is recommended). Installation instructions for different operating systems can be found below:

Windows

- Select *Download R for Windows*
- Select *base*
- Follow the standard instructions for installing programs for Windows. You should choose default start-up options rather than customize start-up options.

Mac OS

- Select *Download R for macOS*

- Look for the desired version of R and click the *.pkg* file to download.
- Open the *.pkg* file and follow the standard instructions for installing applications on MAC OS.
- Drag and drop the R application into the *Applications* folder.

Linux

- Select *Download R for Linux*
- Select the Linux package management system which you are using
- Follow the instructions on the website.

3. R installation

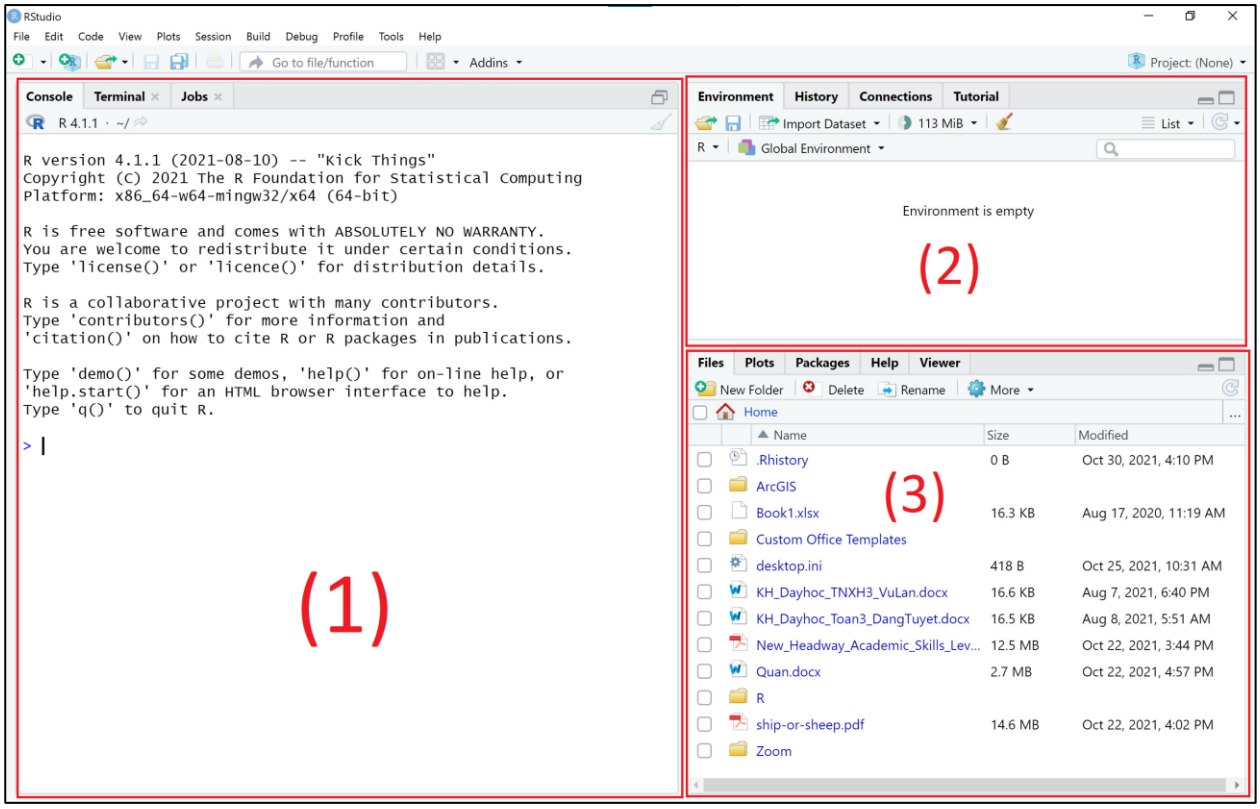
When downloading is finished, you can run the setup program by double clicking the icon, and following the general installation process for every operating system below:

- Select setup program language
- Clicking *Next* to skip the welcome window
- Accept the license (GNU GPL) and click *Next*
- Choose the appropriate location for your program
- Select components to be installed
- Customize the start-up options? -> press *No* (accept defaults) (preference for default setting)
- Select additional tasks
- Finish the installation process

4. R studio

RStudio allows the user to run R in a more user-friendly environment and is available at https://www.rstudio.com/products/rstudio/download/. RStudio is available in open source and commercial editions and runs on the Windows, Mac, and Linux system or in a browser connected to RStudio Server.

When opening RStudio, R is launched as well in the background. You will probably see a layout like this:



(1) The Console tab in RStudio is where you can type, run R code, and see outputs.

(2) The Workspace tab shows all the active objects (variables, imported data, dataframe, etc.) under *Environment*. The *History* tab shows a list of commands used so far.

(3) The *Files* tab shows all the files and folders in your default workspace. The *Plots* tab will show all your graphs. The *Packages* tab will list a series of packages or add-ons needed to run certain processes. For additional information see the *Help* tab.

Console tab

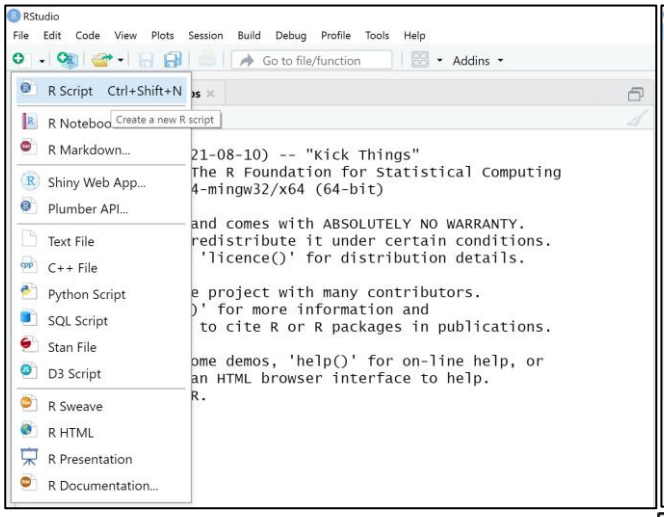
When we open RStudio, the console contains information about the version of R we're working with. Scroll down and try typing a simple expression like 1 + 2. Press the enter key to see the result.

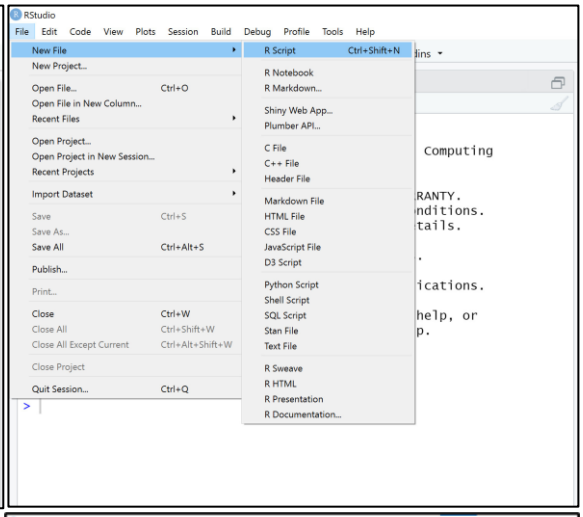
```
> 1 + 2
[1] 3
>
```

As you can see, the console allows you to execute code and see the output immediately. But most of the time, you may wish to save your work with a lot of code lines.

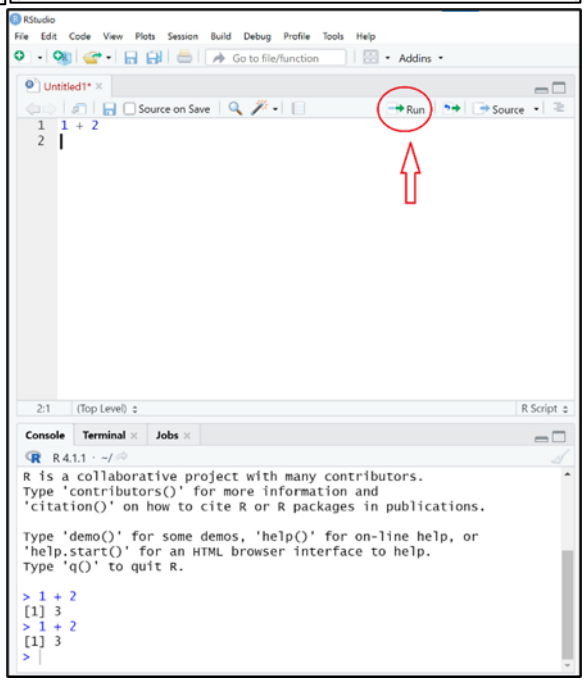
There is a type of document called R-script which can be saved and used later to re-execute the saved code. The script can also be edited so you can execute a modified version of the code.

To create a new R-script you can either go to *File -> New File -> R Script*, or click on the icon with the "+" sign and select "R Script", or simply press Ctrl+Shift+N.



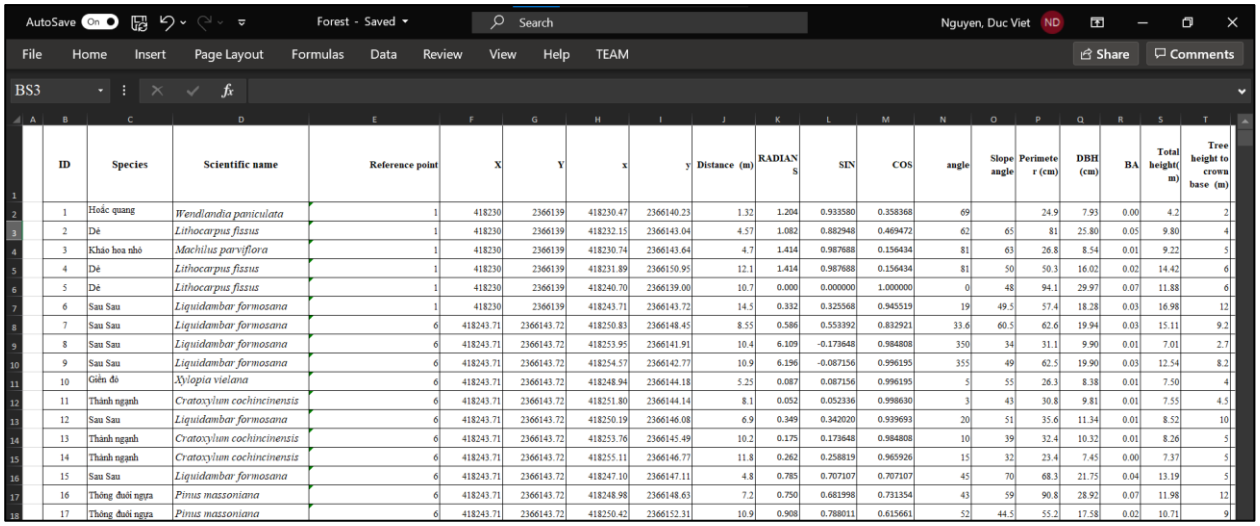


The *Run* is for executing the command (code), or you can leave the cursor on the line where the command is and press Ctrl + Enter. Outputs will appear in the console below.

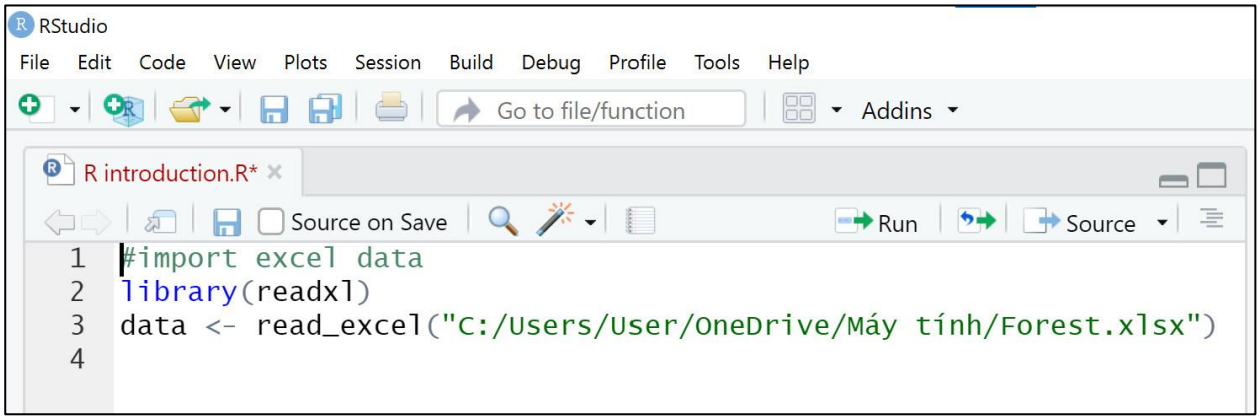


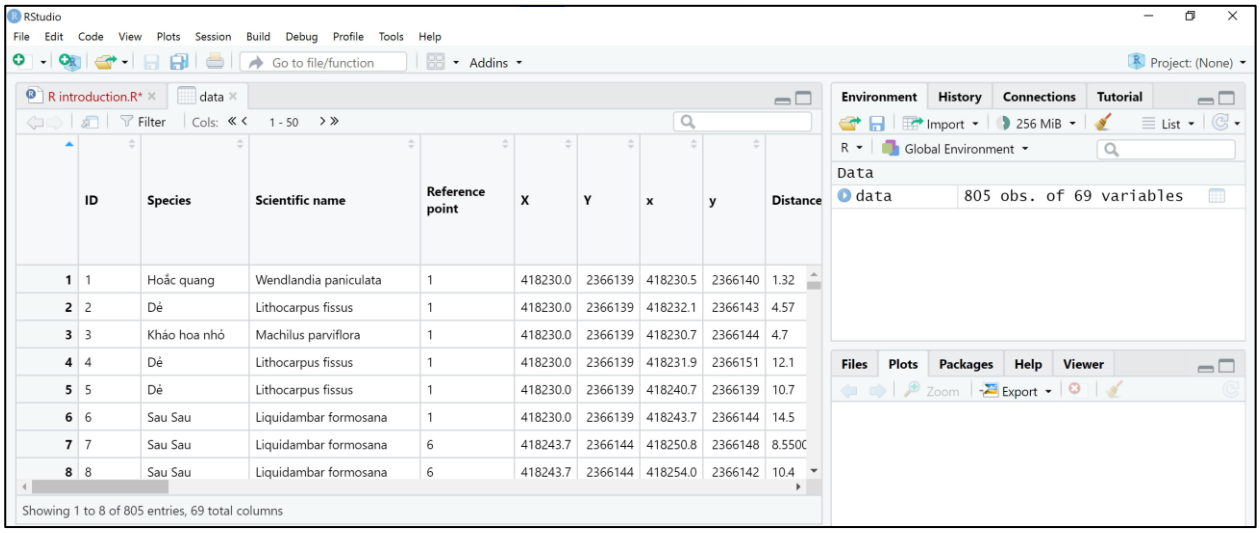
Workspace tab

The workspace tab stores any object, value, dataframe, function, or anything you create during your R session. In the example below, the excel file Forest.xlsx is imported to R and assigned as a dataframe called "data".



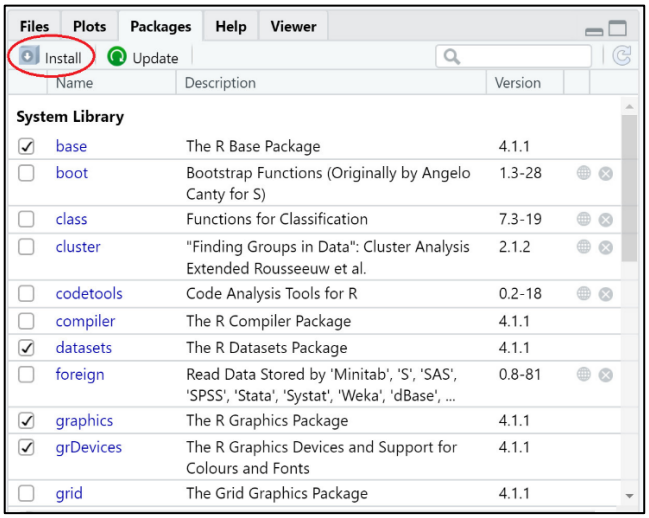
	ID	Species	Scientific name	Reference point	X	Y	x	y	Distance (m)	RADIANS	SIN	COS	angle	Slope angle	Perimeter (cm)	DBH (cm)	BA	Total height(m)	Tree height to crown base (m)
2	1	Hoắc quang	*Wendlandia paniculata*	1	418230	2366139	418230.47	2366140.23	1.32	1.204	0.933580	0.358368	69		24.9	7.93	0.00	4.2	2
3	2	Dẻ	*Lithocarpus fissus*	1	418230	2366139	418232.15	2366143.04	4.57	1.082	0.882948	0.469472	62	65	81	25.80	0.05	9.80	4
4	3	Kháo hoa nhỏ	*Machilus parviflora*	1	418230	2366139	418230.74	2366143.64	4.7	1.414	0.987688	0.156434	81	63	26.8	8.54	0.01	9.22	5
5	4	Dẻ	*Lithocarpus fissus*	1	418230	2366139	418231.89	2366150.95	12.1	1.414	0.987688	0.156434	81	50	50.3	16.02	0.02	14.42	6
6	5	Dẻ	*Lithocarpus fissus*	1	418230	2366139	418240.70	2366139.00	10.7	0.000	0.000000	1.000000	0	48	94.1	29.97	0.07	11.88	6
7	6	Sau Sau	*Liquidambar formosana*	1	418230	2366139	418243.71	2366143.72	14.5	0.332	0.325568	0.945519	19	49.5	57.4	18.28	0.03	16.98	12
8	7	Sau Sau	*Liquidambar formosana*	6	418243.71	2366143.72	418250.83	2366148.45	8.55	0.586	0.553392	0.832921	33.6	60.5	62.6	19.94	0.03	15.11	9.2
9	8	Sau Sau	*Liquidambar formosana*	6	418243.71	2366143.72	418253.95	2366141.91	10.4	6.109	-0.173648	0.984808	350	34	31.1	9.90	0.01	7.01	2.7
10	9	Sau Sau	*Liquidambar formosana*	6	418243.71	2366143.72	418254.57	2366142.77	10.9	6.196	-0.087156	0.996195	355	49	62.5	19.90	0.03	12.54	8.2
11	10	Giền đỏ	*Xylopia vielana*	6	418243.71	2366143.72	418248.94	2366144.18	5.25	0.087	0.087156	0.996195	5	55	26.3	8.38	0.01	7.50	4
12	11	Thành ngạnh	*Cratoxylum cochincinensis*	6	418243.71	2366143.72	418251.80	2366144.14	8.1	0.052	0.052336	0.998630	3	43	30.8	9.81	0.01	7.55	4.5
13	12	Sau Sau	*Liquidambar formosana*	6	418243.71	2366143.72	418250.19	2366146.08	6.9	0.349	0.342020	0.939693	20	51	35.6	11.34	0.01	8.52	10
14	13	Thành ngạnh	*Cratoxylum cochincinensis*	6	418243.71	2366143.72	418253.76	2366145.49	10.2	0.175	0.173648	0.984808	10	39	32.4	10.32	0.01	8.26	5
15	14	Thành ngạnh	*Cratoxylum cochincinensis*	6	418243.71	2366143.72	418255.11	2366146.77	11.8	0.262	0.258819	0.965926	15	32	23.4	7.45	0.00	7.37	5
16	15	Sau Sau	*Liquidambar formosana*	6	418243.71	2366143.72	418247.10	2366147.11	4.8	0.785	0.707107	0.707107	45	70	68.3	21.75	0.04	13.19	5
17	16	Thông đuôi ngựa	*Pinus massoniana*	6	418243.71	2366143.72	418248.98	2366148.63	7.2	0.750	0.681998	0.731354	43	59	90.8	28.92	0.07	11.98	12
18	17	Thông đuôi ngựa	*Pinus massoniana*	6	418243.71	2366143.72	418250.42	2366152.31	10.9	0.908	0.788011	0.615661	52	44.5	55.2	17.58	0.02	10.71	9





	ID	Species	Scientific name	Reference point	X	Y	x	y	Distance
1	1	Hoắc quang	Wendlandia paniculata	1	418230.0	2366139	418230.5	2366140	1.32
2	2	Dẻ	Lithocarpus fissus	1	418230.0	2366139	418232.1	2366143	4.57
3	3	Kháo hoa nhỏ	Machilus parviflora	1	418230.0	2366139	418230.7	2366144	4.7
4	4	Dẻ	Lithocarpus fissus	1	418230.0	2366139	418231.9	2366151	12.1
5	5	Dẻ	Lithocarpus fissus	1	418230.0	2366139	418240.7	2366139	10.7
6	6	Sau Sau	Liquidambar formosana	1	418230.0	2366139	418243.7	2366144	14.5
7	7	Sau Sau	Liquidambar formosana	6	418243.7	2366144	418250.8	2366148	8.550
8	8	Sau Sau	Liquidambar formosana	6	418243.7	2366144	418254.0	2366142	10.4

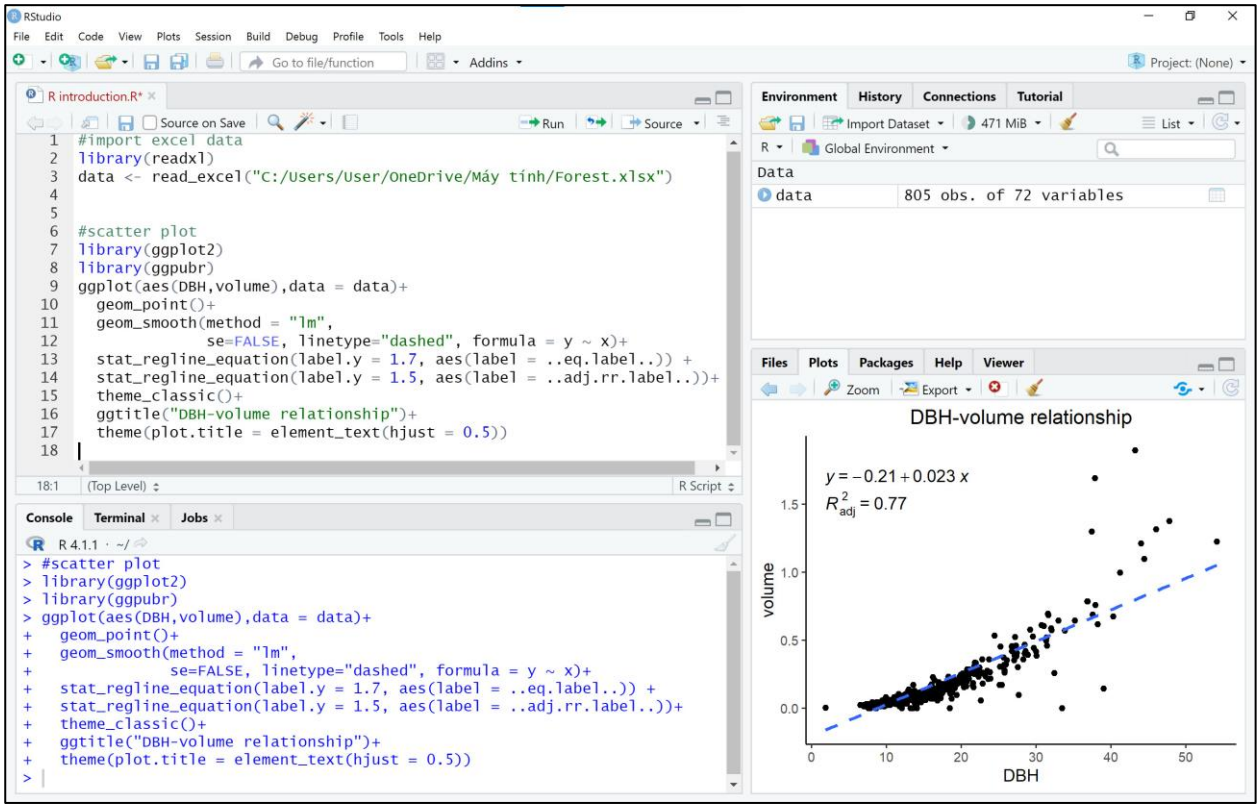
Packages tab



The *package* tab shows the list of packages included in the installation of RStudio. If checked, the package is loaded into R, if not, any command related to that package will not work. You can also install other packages by clicking on the '*Install*' icon. Another way to install a package is by typing the command *install.packages*("*packages name*"). For example, to install package "ggplot2", the code is: `install.packages("ggplot2")`.

Plots tab

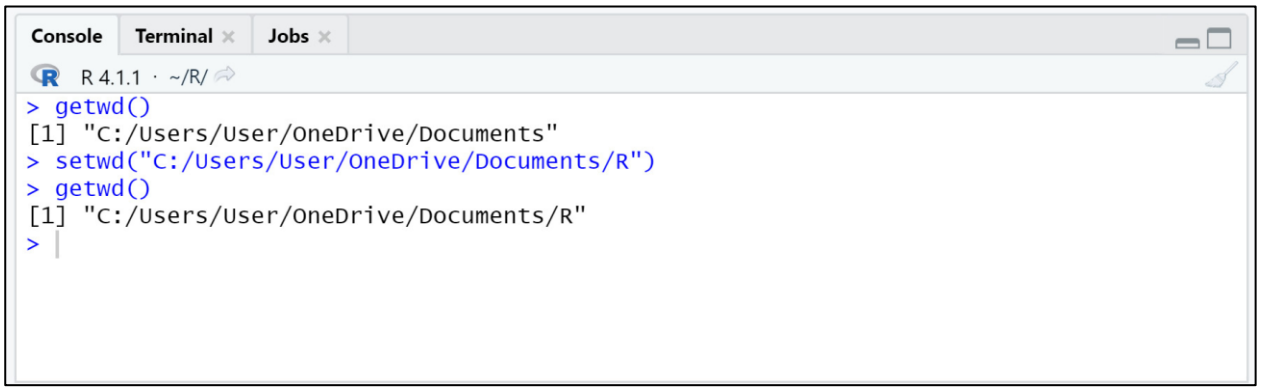
The plot tab will display all the graphs created during your R session. The below example shows the scatter plot generated by the command on line 9.



Working directory

The working directory is a folder on your computer that is set as the default location where R reads and saves files.

For showing the working directory, use the command `getwd()`. The working directory can be changed by the command `setwd()`.



5. References

Irizarry, R. A. (2021). *Chapter 1 Getting started with R and RStudio | Introduction to Data Science*. https://rafalab.github.io/dsbook/getting-started.html