



Food and Agriculture
Organization of the
United Nations

SUSTAINABLE
DEVELOPMENT
GOALS

Introduction to R and RStudio



**2022 Virtual Training on Data Disaggregation and Small Area Estimation for the SDGs –
22-25 November 2022, INStAD Benin, INSTAT Mali, Stats South Africa, Statistics Botswana**

General overview

- **R** is a programming language for statistical computing, data analytics and scientific research. It is one of the most widely used languages by statisticians, data analysts and researchers to manage, manipulate, analyze and visualize data.
- **Rstudio** is an integrated development environment for R that allows users to interact more easily with R by integrating different aspects of scripting, from code completion to debugging.

Note: On July 27, 2022, it was announced that Rstudio was changing its name to Posit to move from an R-exclusive tool to a language agnostic system.

In order to use RStudio, R needs to be installed first.



R and RStudio Installation

- **Step 1:** Go to the [Comprehensive R Archive Network \(CRAN\)](#) website and click on “Download R for Windows” (or “Download R for MacOS”).
- **Step 2:** Click on the “base” subdirectory link (or the package link, *.pkg file*).
- **Step 3:** Click on “Download R-4.2.1 for Windows” (the R version available might be different depending on updates produced after the compilation of this ppt). The link allows downloading an installer extension (*.exe file*).
- **Step 4:** Run the *.exe* file and step through the installation wizard accepting the default settings.

Once R is installed, you can then proceed to the installation of RStudio.

- **Step 5:** Go to [RStudio download](#) website and click on the “Download RStudio for Windows” button (or the link for the MacOS version).
- **Step 6:** Run the *.exe* file and follow the installation instructions.



RStudio interface

The image shows the RStudio interface with two callouts. The first callout, a blue oval, is positioned over the Command Console and contains the text "Command Console" in red. The second callout, also a blue oval, is positioned over the Environment/History/Connections/Tutorial panes and contains the text "Environment/History/Connections/Tutorial panes" in red. The RStudio interface includes a menu bar (File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help), a toolbar, a console window with R version information and a prompt, and a right-hand pane with tabs for Environment, History, Connections, and Tutorial. The Environment pane shows "Global Environment" and "Environment is empty". Below the Environment pane are tabs for Files, Plots, Packages, Help, and Viewer.

```
R version 4.0.3 (2020-10-10) -- "Bunny-wunnies Freak Out"  
Copyright (C) 2020 The R Foundation for Statistical Computing  
Platform: x86_64-w64-mingw32/x64 (64-bit)  
  
R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.  
  
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
> |
```

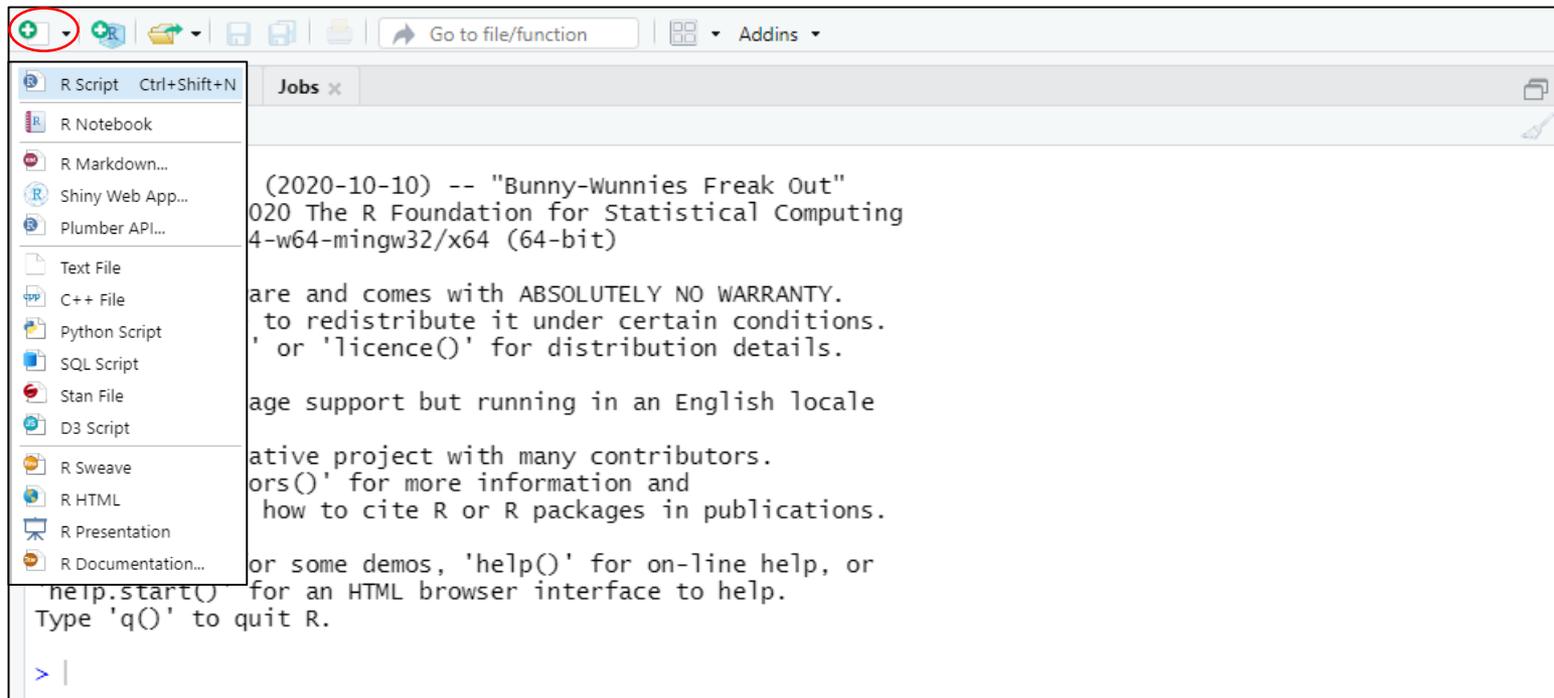
Environment History Connections Tutorial
Global Environment
Environment is empty
Files Plots Packages Help Viewer
Zoom Export

Command Console

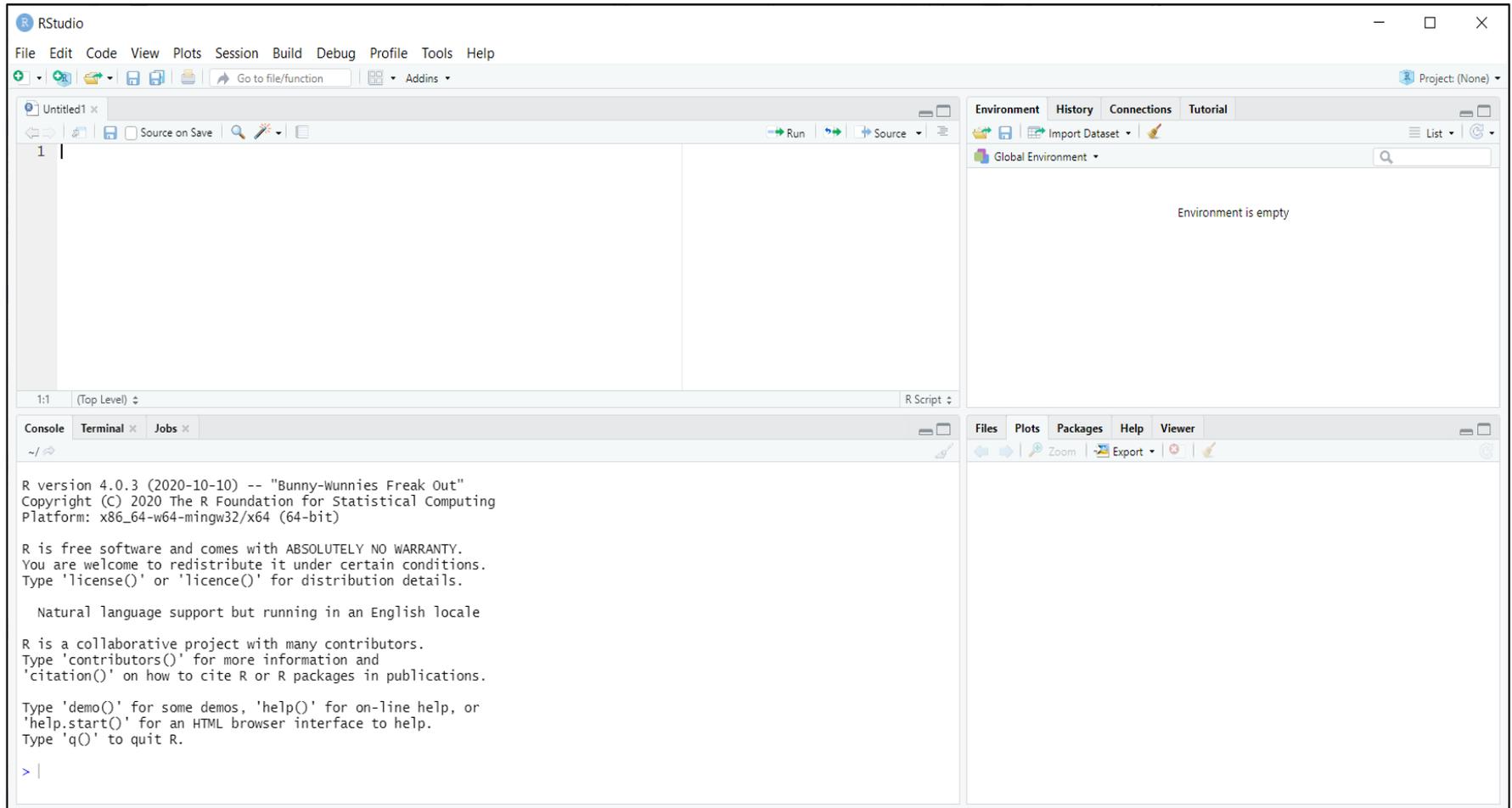
Environment/History/
Connections/Tutorial
panes

Files/Plots/Packages
/Help/Viewer panes

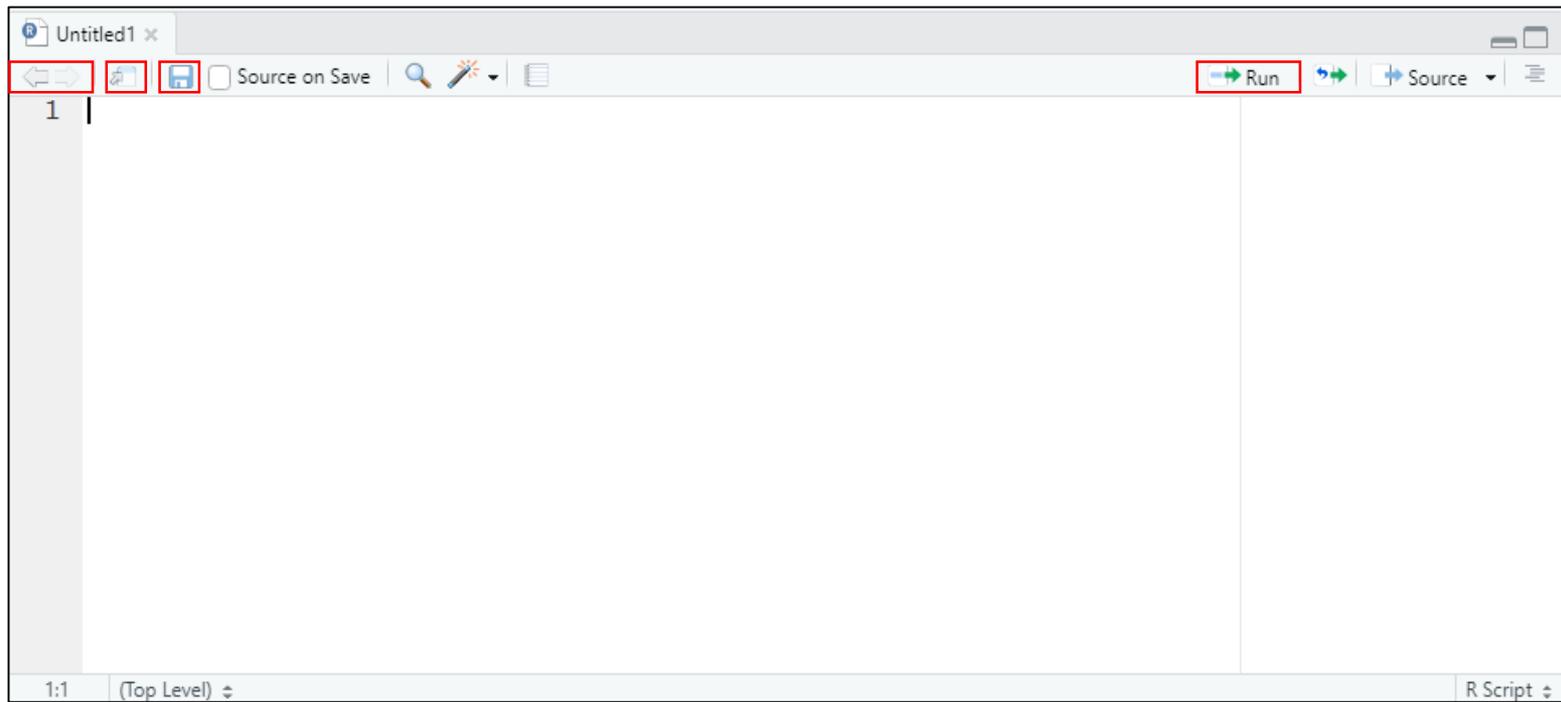
RStudio: command console



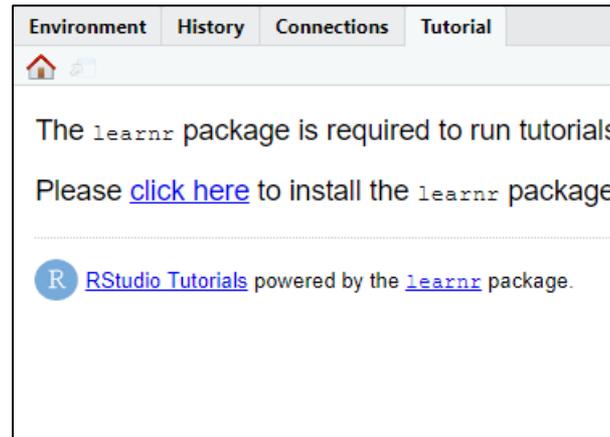
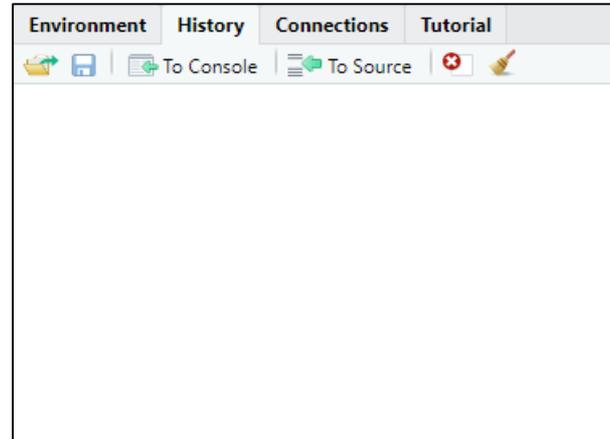
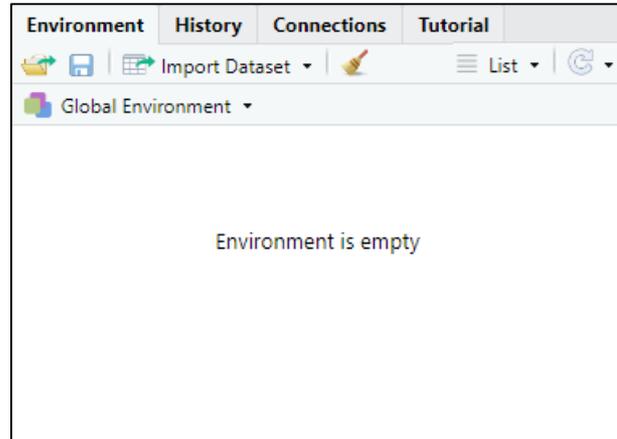
RStudio interface with editor window open



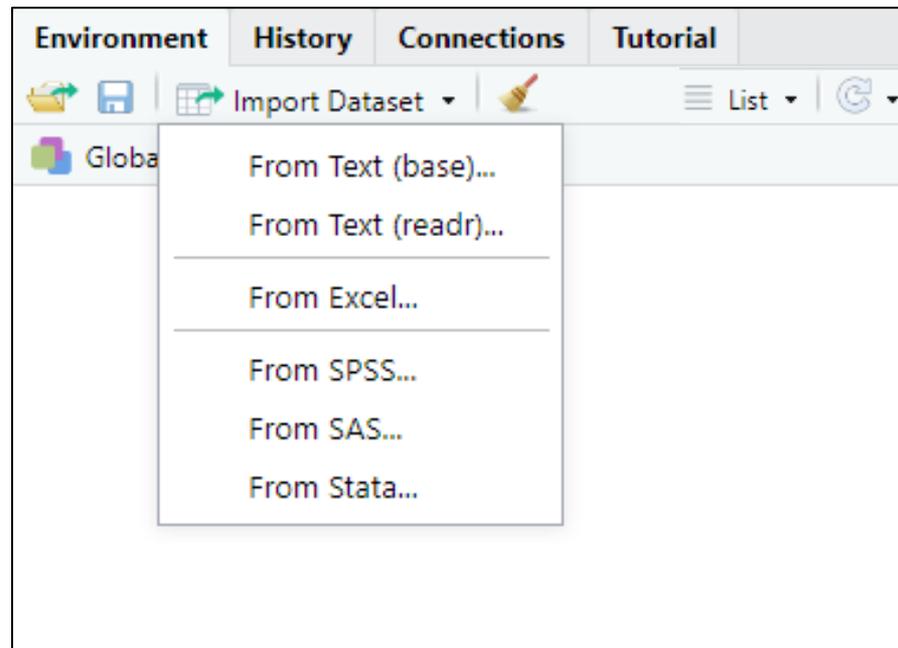
RStudio: editor window



RStudio: Environment/History/Connections/Tutorial panels



Import dataset tab



Import dataset tab (2)

Import Excel Data

File/URL: Browse...

Data Preview:

Sepal.Length <small>(double)</small>	Sepal.Width <small>(double)</small>	Petal.Length <small>(double)</small>	Petal.Width <small>(double)</small>	Species <small>(character)</small>
5.1	3.5	1.4	0.2	se
4.9	3.0	1.4	0.2	se
4.7	3.2	1.3	0.2	se
4.6	3.1	1.5	0.2	se
5.0	3.6	1.4	0.2	se
5.4	3.9	1.7	0.4	se
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa

Previewing first 50 entries.

Import Options:

Name: Max Rows: First Row as Names

Sheet: Skip: Open Data Viewer

Range: NA:

Code Preview:

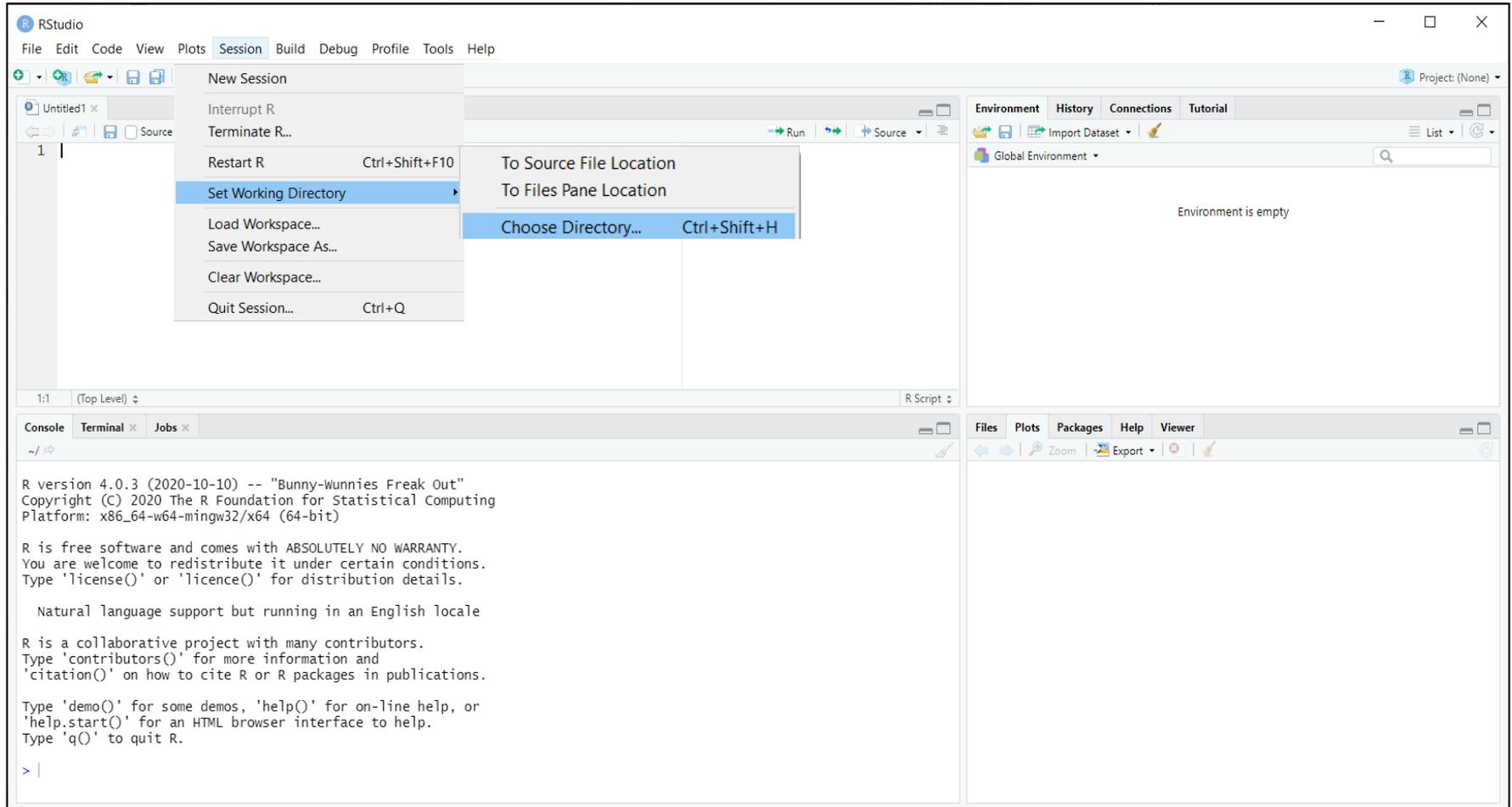
```
library(readxl)
iris <- read_excel("iris.xlsx")
View(iris)
```

[? Reading Excel files using readxl](#)

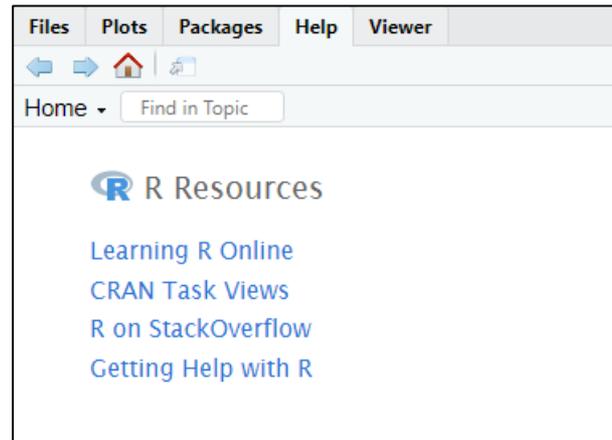
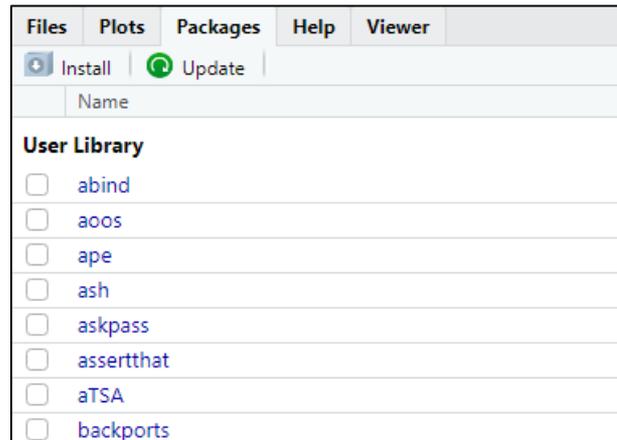
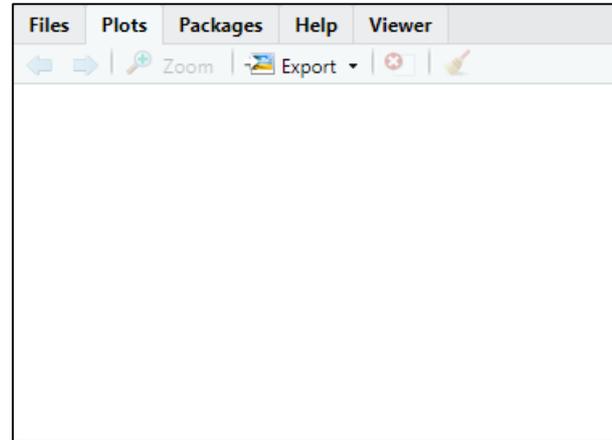
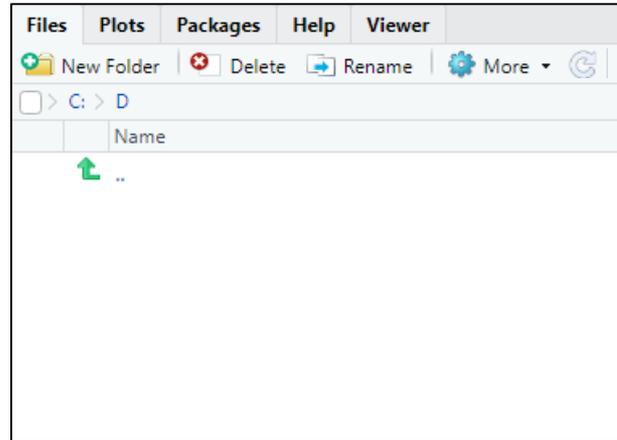
Import Cancel



Setting the working directory



RStudio: Files/Plots/Packages/Help/Viewer panels



Packages

- **Packages:** extensions that contain code, data, and documentation in a standardized format that can be installed and used by users of R to solve specific analytical problems.
- The base version of R already includes many useful packages that allow performing elementary tasks such as simple calculations, data exploration, and loading of text data files.
- Complex tasks may require intensive coding using base packages functions, but are made easier by user-supplied packages.



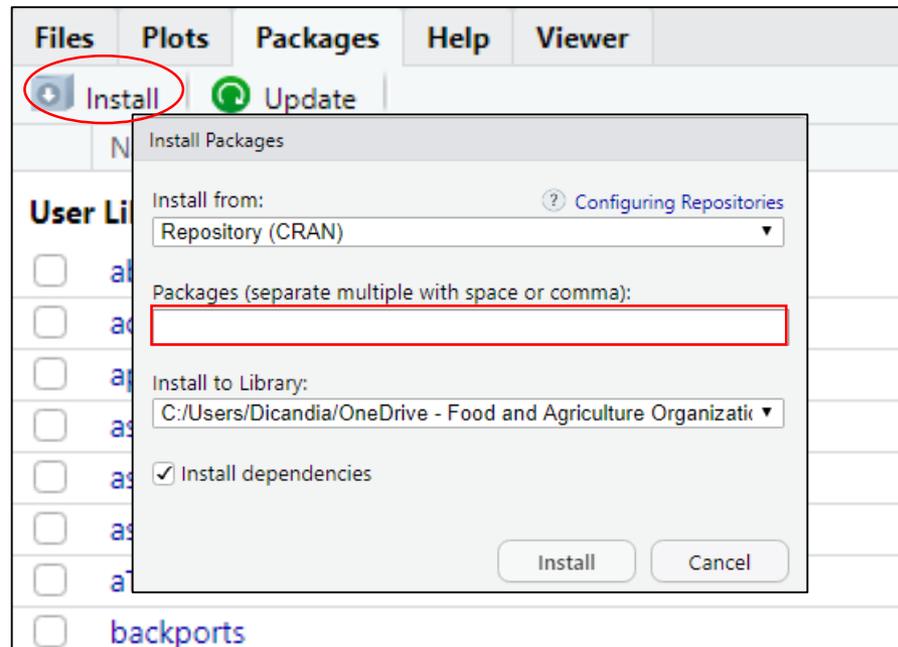
Packages (2)

To **install a package** from CRAN (the official repository for user contributed R packages) and then load it, use the following commands:

```
install.packages("name of the package")
```

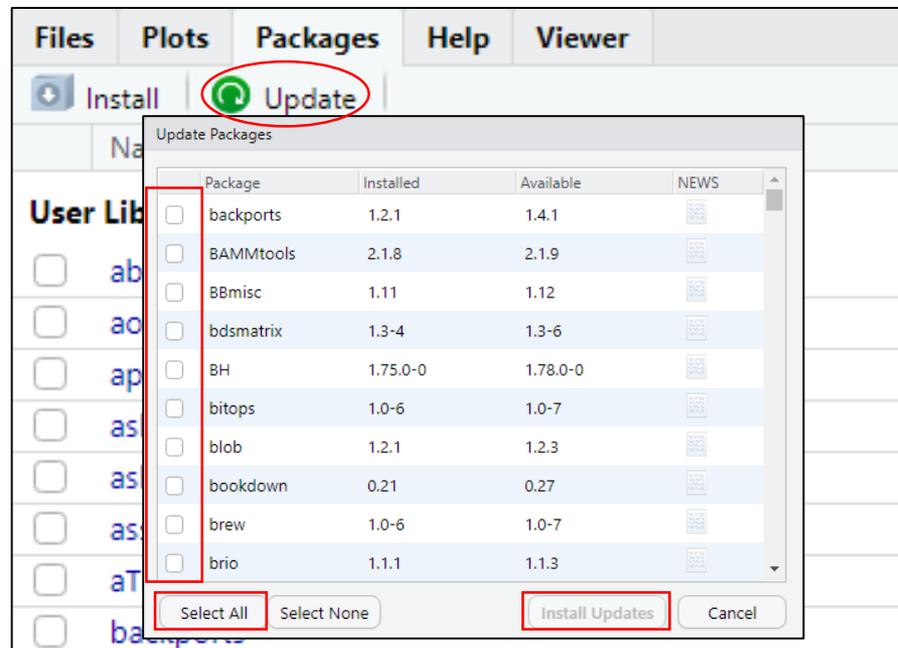
```
library("name of the package")
```

Another way to install packages is to use the Install Packages tab:



Packages (3)

- R packages are sometimes **updated** to improve or modify functionality. It is advisable to occasionally update the packages installed on your computer.
- You can update your installed R packages in RStudio by clicking the Update button in the toolbar in the Packages panel.



List of packages used in the training

For importing	Description	For data manipulation	Description
openxlsx	<i>Read, Write and Edit xlsx Files</i>	dplyr	<i>A fast, consistent tool for working with data frame</i>
readxl	<i>Only read excel files</i>	tidyr	<i>Tools to help to create tidy data</i>
data.table	<i>Fast file reader, writer</i>		
haven	<i>Import and Export 'SPSS', 'Stata' and 'SAS' Files</i>		

For various uses	Description
Boruta	<i>Wrapper algorithm for all relevant Feature Selection</i>
relaimpo	<i>Relative Importance of regressors in linear models</i>
car	<i>Companion to Applied Regression</i>
lme4	<i>Fit linear and generalized linear mixed-effects models</i>
RLRsim	<i>Exact (Restricted) Likelihood Ratio Tests for Mixed and Additive Models</i>
RColorBrewer	<i>Provides color schemes for graphics</i>

List of packages used in the training

For analysis of complex survey samples
survey
ReGenesees

For Small Area Estimation	sae	emdi	mind	SAEval
----------------------------------	------------	-------------	-------------	---------------

For data visualization	ggplot2
-------------------------------	----------------

For spatial data	Description
raster	<i>Reading, writing, manipulating, analyzing and modeling of spatial data.</i>
tabularaster	<i>Tidy Tools for 'Raster' Data</i>
sp	<i>Classes and methods for spatial data</i>
rgdal	<i>Bindings for the 'Geospatial' Data Abstraction Library</i>
sf	<i>Support for simple features, a standardized way to encode spatial vector data</i>
spdep	<i>Provide tools to study spatial dependence</i>
geodata	<i>Functions for downloading of geographic data for use in spatial analysis and mapping.</i>

PRACTICAL EXAMPLES IN RSTUDIO



Thank you!

