

Minería de Textos y Análisis de Sentimientos

Curso de Formación Continua | Universidad de Córdoba | Curso 2025/26



Fundamentos básicos en R

Jose V. Die Ramón

Córdoba, octubre 2025

Módulo : Fundamentos básicos de R

- Instalación de R & RStudio
- Instalación de paquetes en R
- Ejecución de scripts en RStudio
- Tipos de objetos : vectores y tablas de datos
- Manipulación de objetos

R

- **INTERACTIVO**, lenguaje de programación. Deriva del paquete estadístico S. Analistas de Datos
- **EJECUTABLE**, como en cualquier otro lenguaje de programación podemos crear scripts
- **LIBRE** y **CÓDIGO ABIERTO**
- **MULTIPLATAFORMA**, Windows, Mac OS, Linux, Unix
- **COMUNIDAD de USUARIOS**, Stackoverflow

INSTALACIÓN DE SOFTWARE

- **R** , lenguaje de programación que usaremos para el Análisis de Datos / Textos / Sentimiento
- **RStudio**, entorno en formato escritorio que usaremos para editar, organizar y probar nuestros scripts

INSTALACIÓN DE SOFTWARE

- **R , Comprehensive R Archive Network (CRAN)**

The screenshot shows a Google search results page for the query "CRAN". The search bar at the top contains "CRAN". Below the search bar, there are navigation links for "Todo", "Imágenes", "Maps", "Noticias", "Vídeos", "Más", "Configuración", and "Herramientas". A message indicates approximately 17,600,000 results found in 0.50 seconds. The first result is titled "The Comprehensive R Archive Network" and includes a link to <https://cran.r-project.org/>. Below the title, there is a snippet of text: "What are R and CRAN? R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of ...". To the right of this result, there are two additional sections: "Contributed Packages" and "CRAN Packages By Name". The "Contributed Packages" section has a link to "Más resultados de r-project.org ». The "CRAN Packages By Name" section lists letters from A to L followed by "...". At the bottom of the screenshot, there is a separate entry for "Cran - R Project" with a similar structure.

Aproximadamente 17.600.000 resultados (0,50 segundos)

The Comprehensive R Archive Network
<https://cran.r-project.org/> ▾ Traducir esta página
What are R and CRAN? R is 'GNU S', a freely available language and environment for statistical computing and graphics which provides a wide variety of ...

Contributed Packages
Contributed Packages. Available Packages. Currently, the CRAN ...
[Más resultados de r-project.org »](#)

CRAN Packages By Name
Available CRAN Packages By Name.
A B C D E F G H I J K L ...

Cran - R Project
<https://cran-archive.r-project.org/> ▾ Traducir esta página
[DIR], bin/, 2016-02-27 10:54, -. [DIR], web/, 2018-02-22 19:16, -. Apache/2.4.10 (Debian) Server at cran-archive.r-project.org Port 443.

INSTALACIÓN DE SOFTWARE

- **R , Comprehensive R Archive Network (CRAN)**



CRAN
[Mirrors](#)
[What's new?](#)
[Search](#)
[CRAN Team](#)

About R
[R Homepage](#)
[The R Journal](#)

Software
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Task Views](#)
[Other](#)

Documentation
[Manuals](#)
[FAQs](#)

Donations
[Donate](#)

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.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2025-06-13, Great Square Root) [R-4.5.1.tar.gz](#), read [what's new](#) in the latest version.
- The CRAN directory [src/base-prerelease](#) contains R alpha, beta, and rc releases as daily snapshots in time periods before a planned release.
- Between releases, the same directory [src/base-prerelease](#) contains snapshots of current patched and development versions. Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Alternatively, daily snapshots are [available here](#).
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#).

INSTALACIÓN DE SOFTWARE

- **R**, Comprehensive R Archive Network (**CRAN**)

R for Windows

Subdirectories:

[base](#) Binaries for base distribution. This is what you want to [install R for the first time](#).

[contrib](#) Binaries of contributed CRAN packages (for R >= 4.0.x).

[old contrib](#) Binaries of contributed CRAN packages for outdated versions of R (for R < 4.0.x).

[Rtools](#) Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

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.

subdirectorio **base**

INSTALACIÓN DE SOFTWARE

- **R , Comprehensive R Archive Network (CRAN)**

R-4.5.1 for Windows

[Download R-4.5.1 for Windows](#) (86 megabytes, 64 bit)

[README on the Windows binary distribution](#)

[New features in this version](#)

This build requires UCRT, which is part of Windows since Windows 10 and Windows Server 2016. On older systems, UCRT has to be installed manually from [here](#).

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.

Frequently asked questions

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

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

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

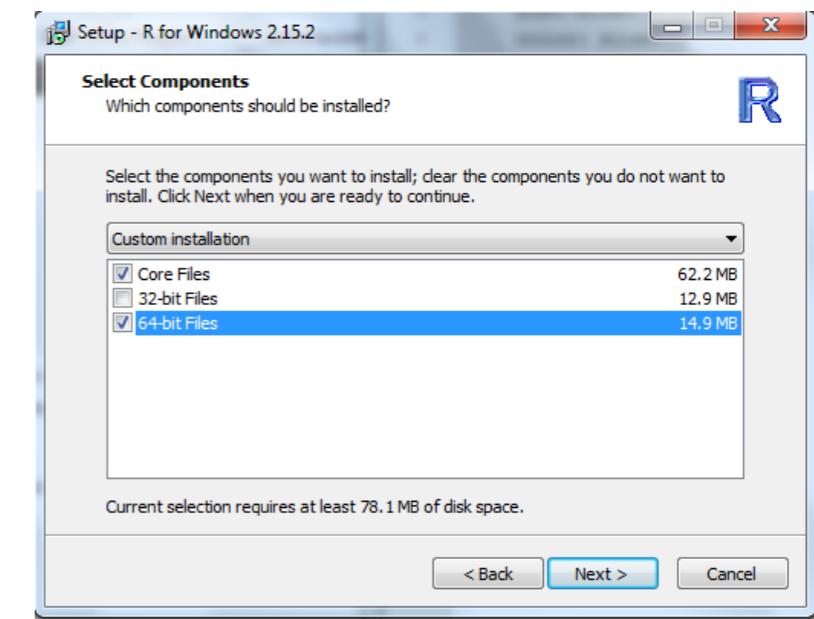
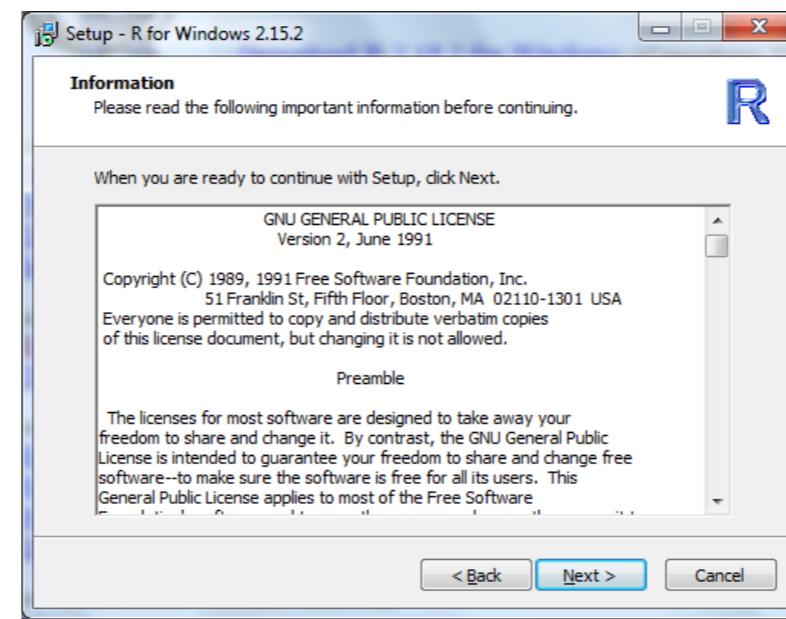
Note to webmasters: A stable link which will redirect to the current Windows binary release is
<<CRAN MIRROR>/bin/windows/base/release.html>.

Last change: 2025-06-13

subdirectorio **base**

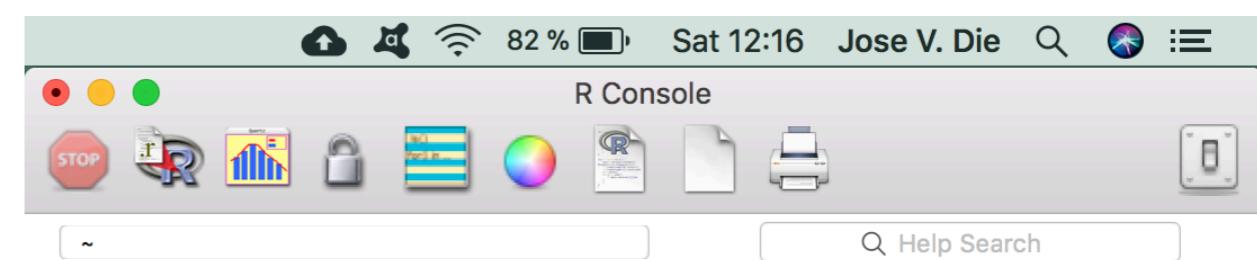
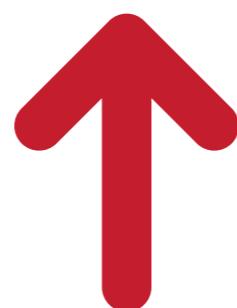
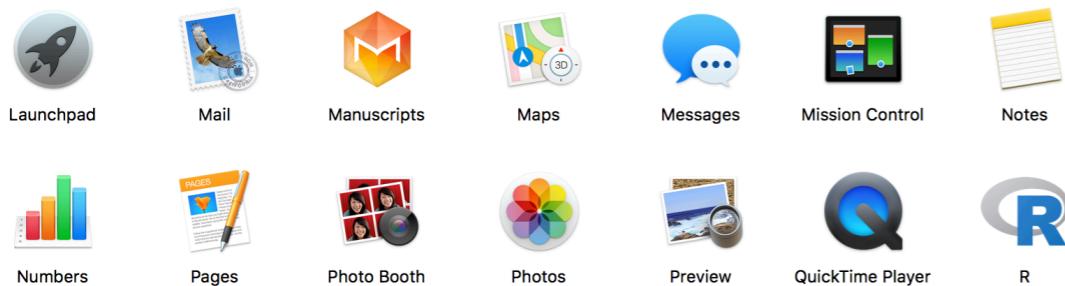
INSTALACIÓN DE SOFTWARE

- **R , Comprehensive R Archive Network (CRAN)**



INSTALACIÓN DE SOFTWARE

- **R , Comprehensive R Archive Network (CRAN)**



R version 3.5.0 (2018-04-23) -- "Joy in Playing"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (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.

[R.app GUI 1.70 (7521) x86_64-apple-darwin15.6.0]

>

INSTALACIÓN DE SOFTWARE

- **RStudio**, entorno en formato escritorio que usaremos para editar, organizar y probar nuestros scripts

The screenshot shows a Google search results page. The search query 'rstudio download' is entered in the search bar. Below the search bar, there is a navigation menu with options: Todo, Vídeos, Imágenes, Vídeos cortos, Noticias, Web, Libros, Más ▾, and Herramientas ▾. The 'Todo' option is underlined, indicating it is the active category.

Sugerencia: Mostrar resultados en **español**. También puedes consultar más información sobre [cómo filtrar por idioma](#).

 Posit
<https://posit.co/download/rstudio> · Traducir esta página

RStudio Desktop

The **RStudio integrated development environment (IDE)** is a set of tools built to help you be more productive with R and Python.

 Posit
<https://posit.co/downloads> · Traducir esta página

Download RStudio

The **RStudio integrated development environment (IDE)** is a set of tools built to help you be more productive with R and Python.

INSTALACIÓN DE SOFTWARE

- **RStudio** , entorno en formato escritorio que usaremos para editar, organizar y probar nuestros scripts

1: Install R

RStudio requires R 3.6.0+. Choose a version of R that matches your computer's operating system.

R is not a Posit product. By clicking on the link below to download and install R, you are leaving the Posit website. Posit disclaims any obligations and all liability with respect to R and the R website.

[DOWNLOAD AND INSTALL R](#)

2: Install RStudio

Choose macOS version

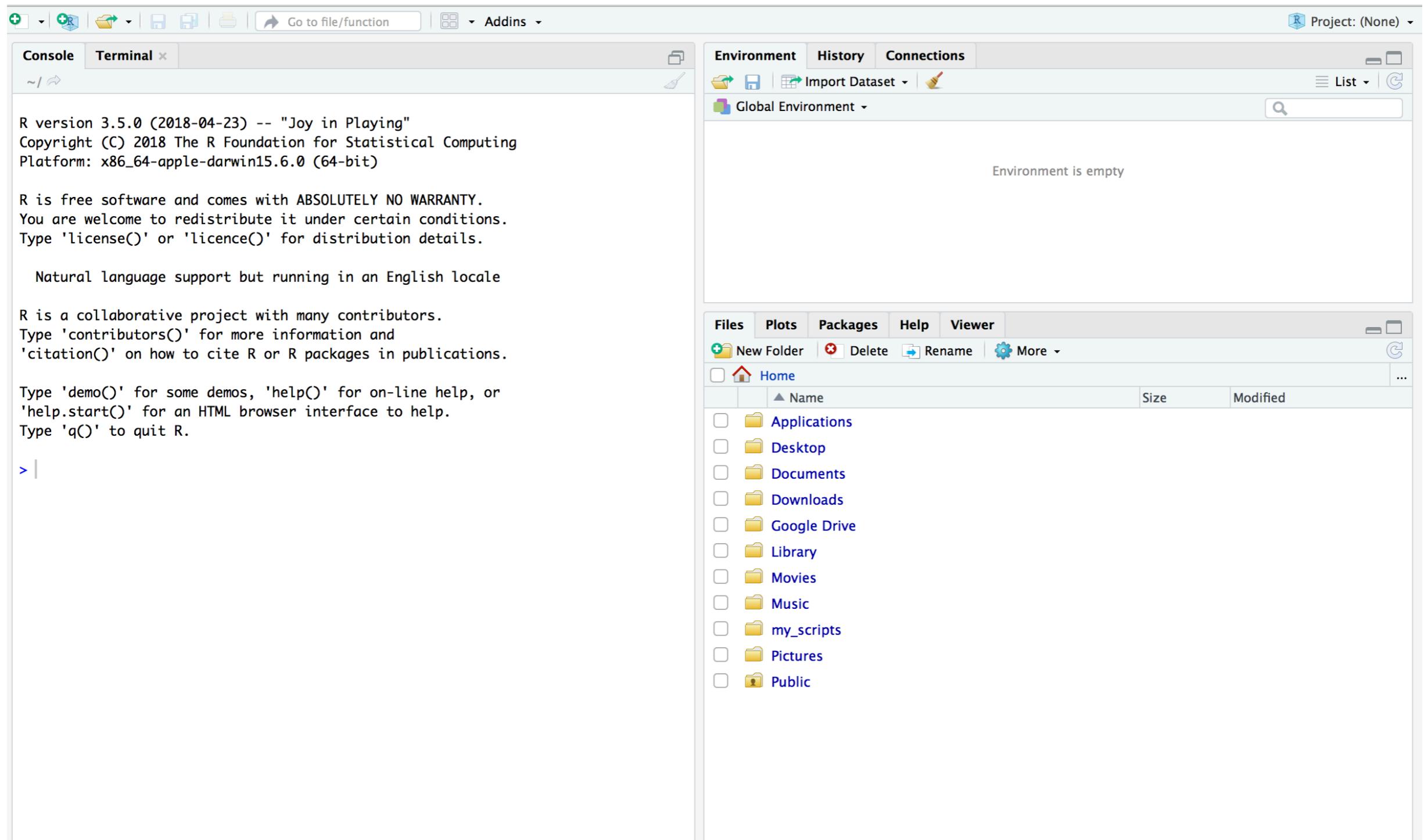
The current version of RStudio is only supported on **macOS 13 and higher**. Please select your macOS version to get the correct link.

INSTALACIÓN DE SOFTWARE

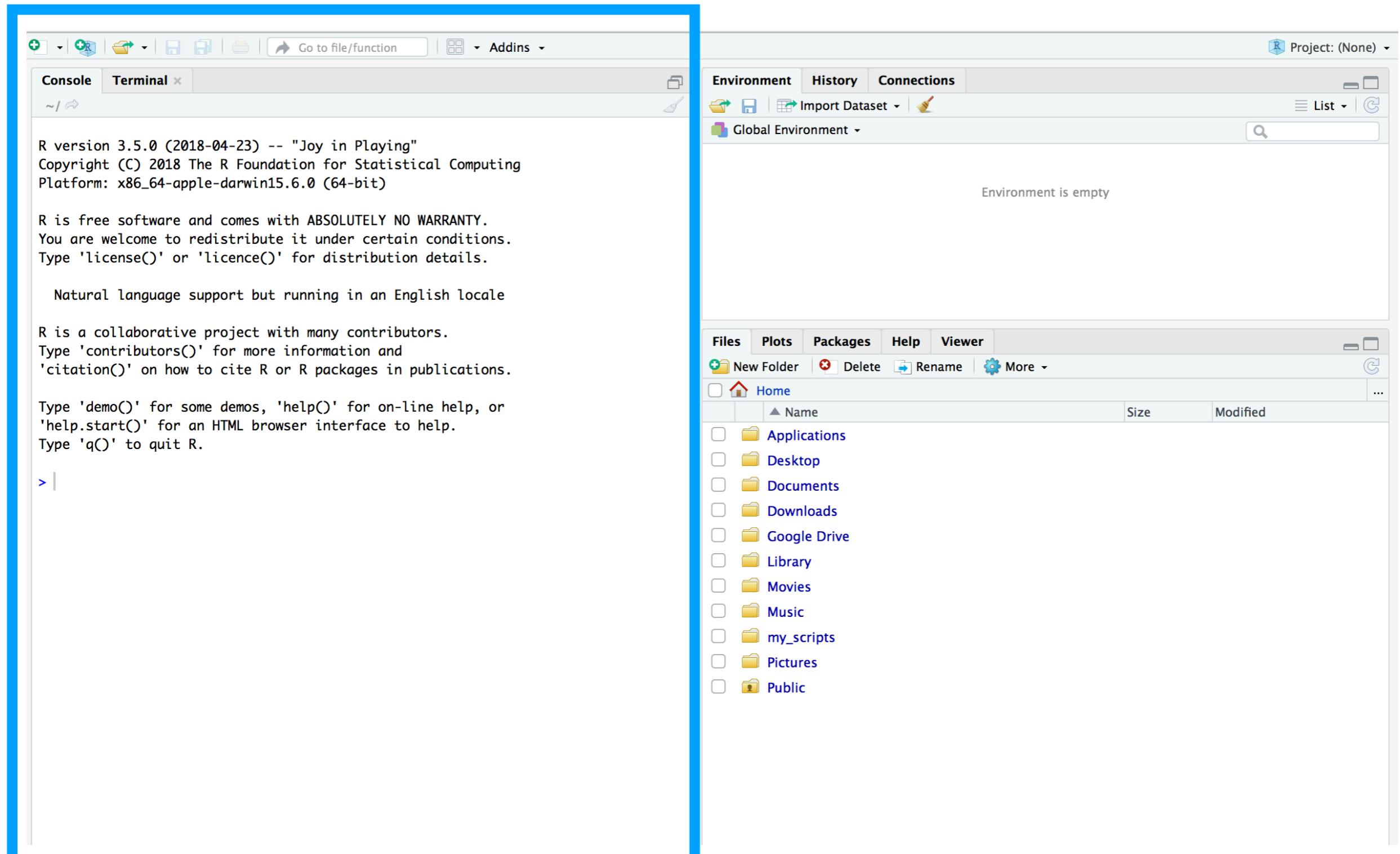
- **RStudio** , entorno en formato escritorio que usaremos para editar, organizar y probar nuestros scripts

OS	Download	Size	SHA-256
Windows 10/11	RSTUDIO-2025.09.0-387.EXE ↓	287.97 MB	8CE88C63
macOS 13+	RSTUDIO-2025.09.0-387.DMG ↓	634.56 MB	8568D611
Ubuntu 22/Debian 12	RSTUDIO-2025.09.0-387-AMD64.DEB ↓	216.76 MB	851FB642
Ubuntu 24	RSTUDIO-2025.09.0-387-AMD64.DEB ↓	216.76 MB	851FB642

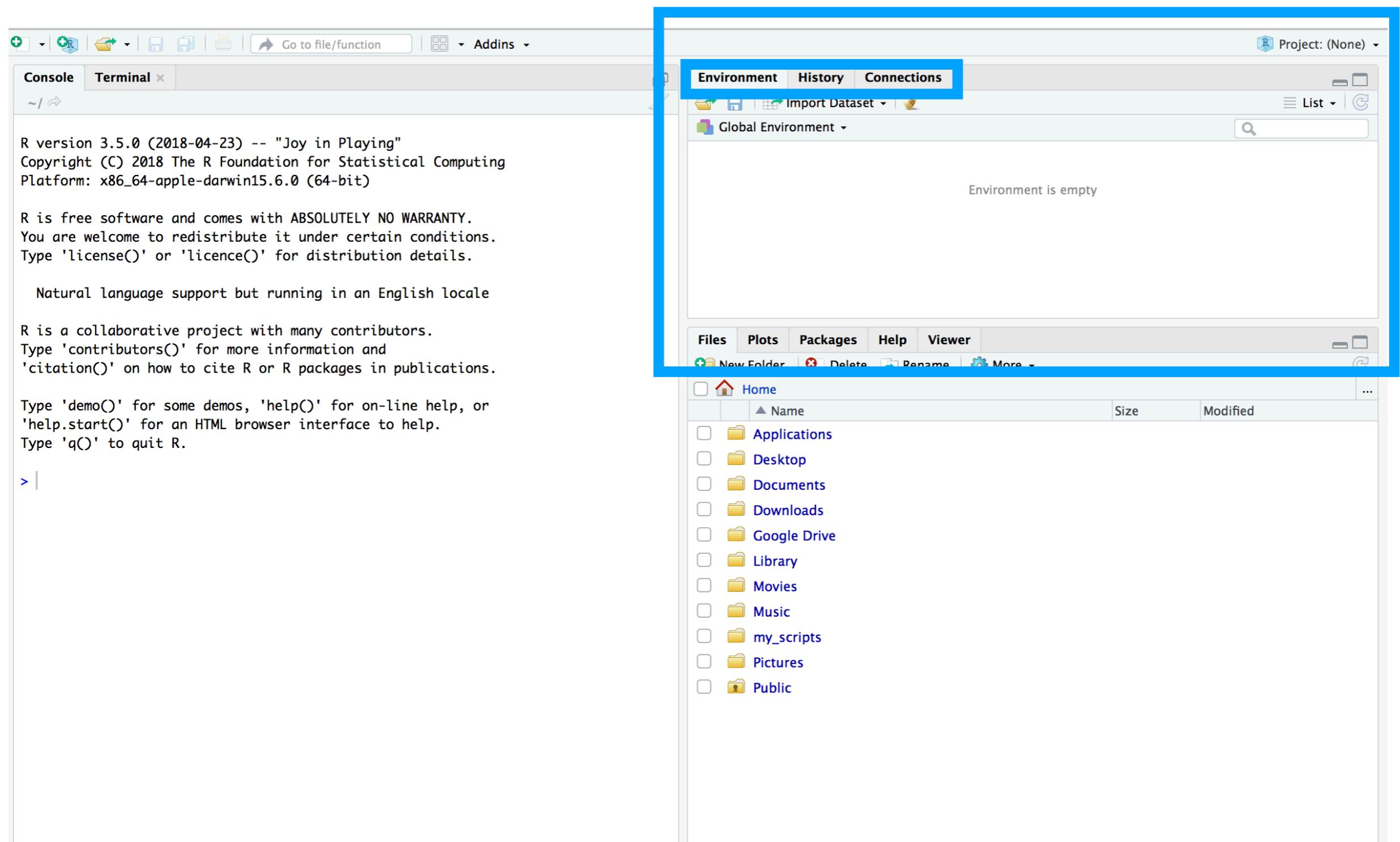
USANDO RSTUDIO POR PRIMERA VEZ



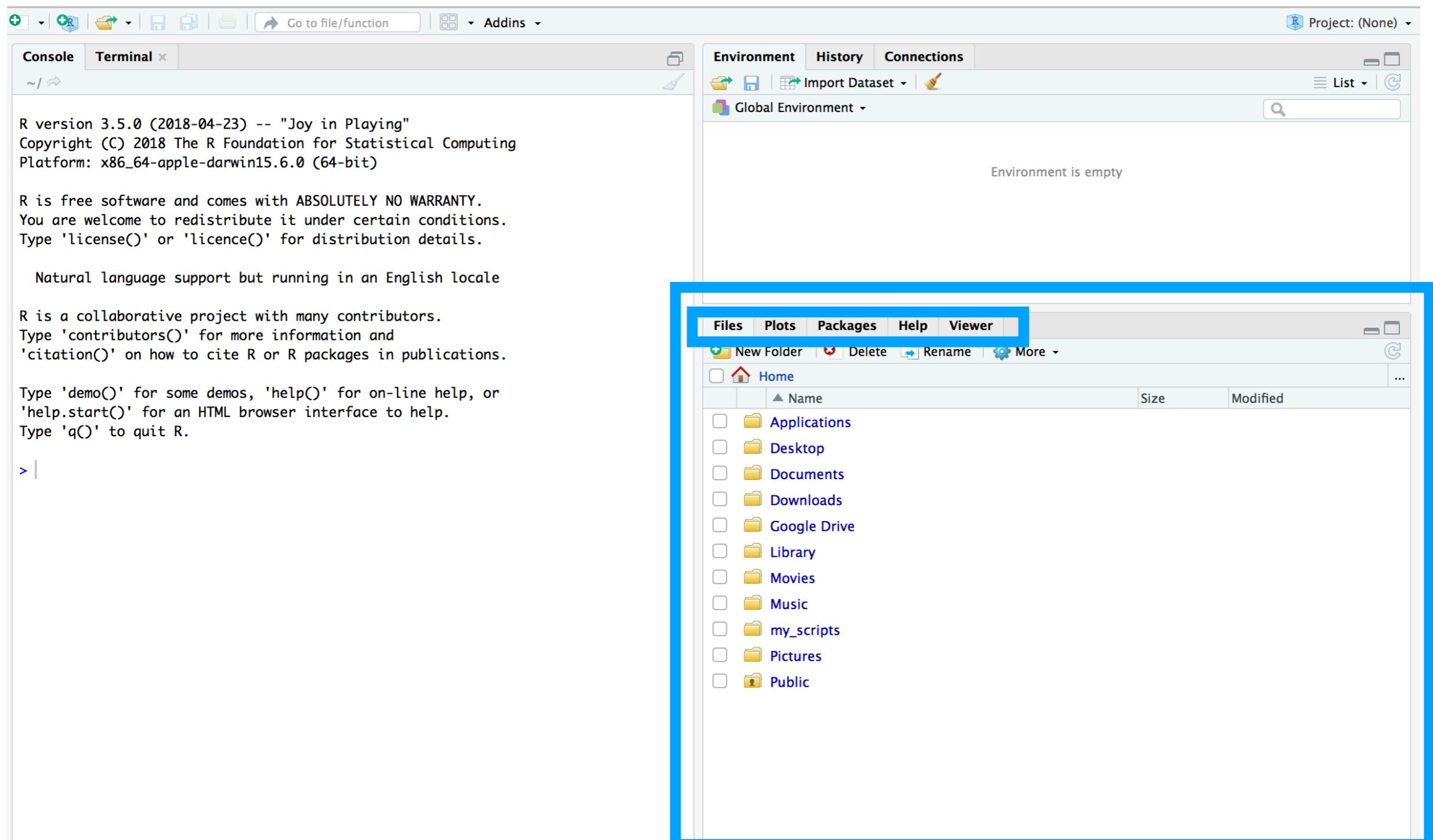
USANDO RSTUDIO POR PRIMERA VEZ



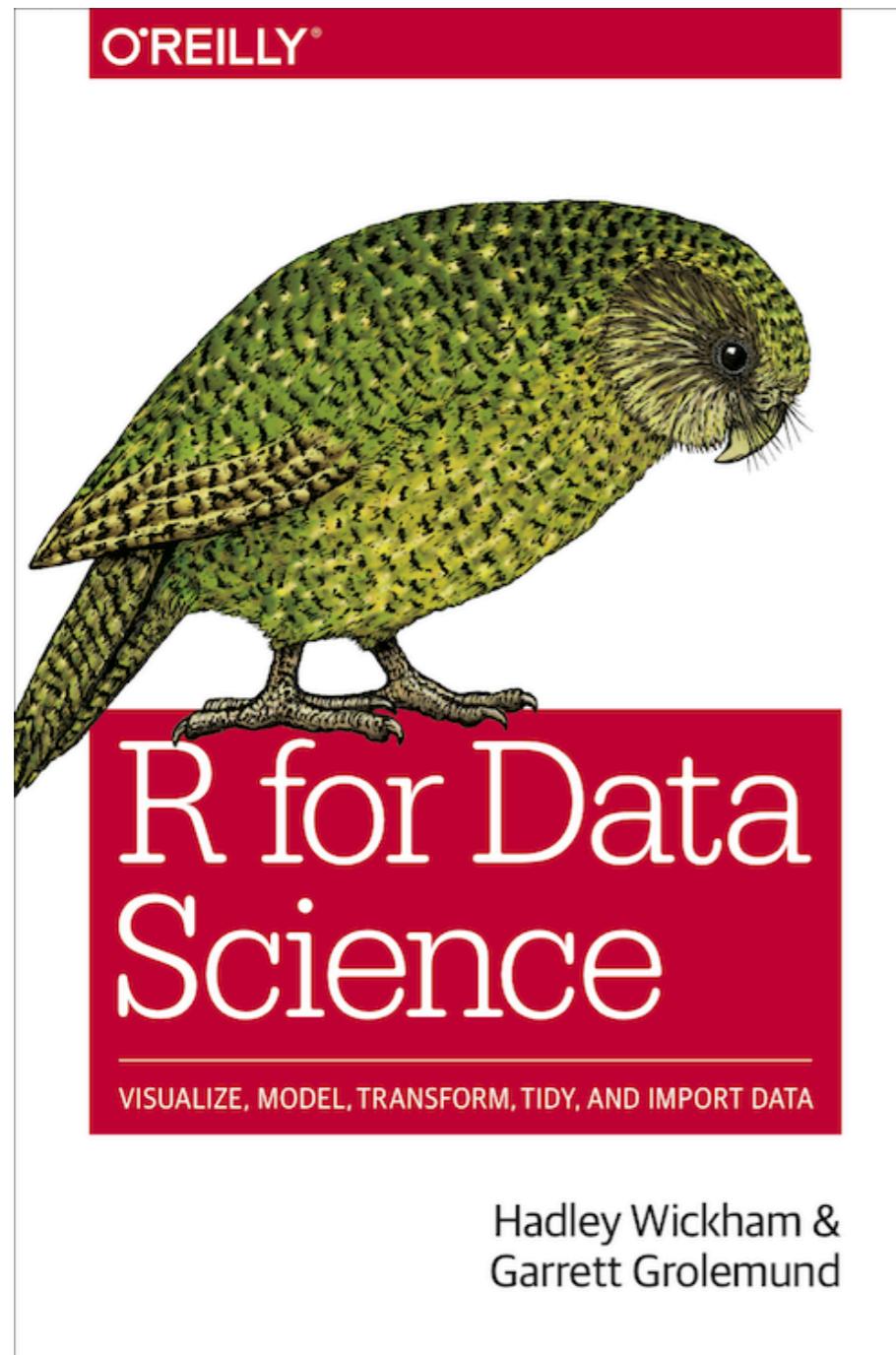
USANDO RSTUDIO POR PRIMERA VEZ



USANDO RSTUDIO POR PRIMERA VEZ



OPCIONES GLOBALES



- **What is real ?**

As a beginning R user, it's OK to consider your environment (i.e. the objects listed in the environment pane) "real". However, in the long run, you'll be much better off if you **consider your R scripts as "real"**.

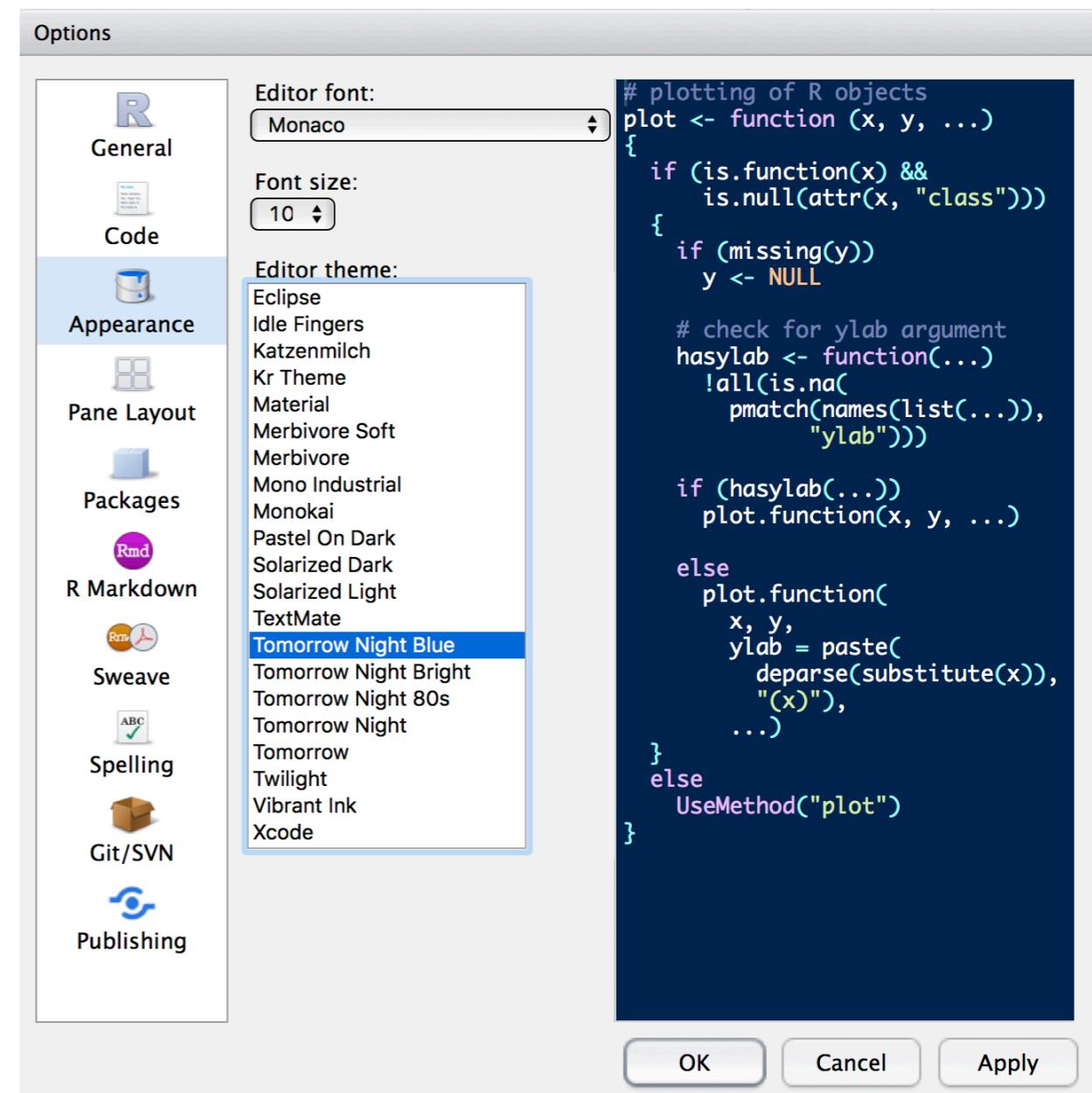
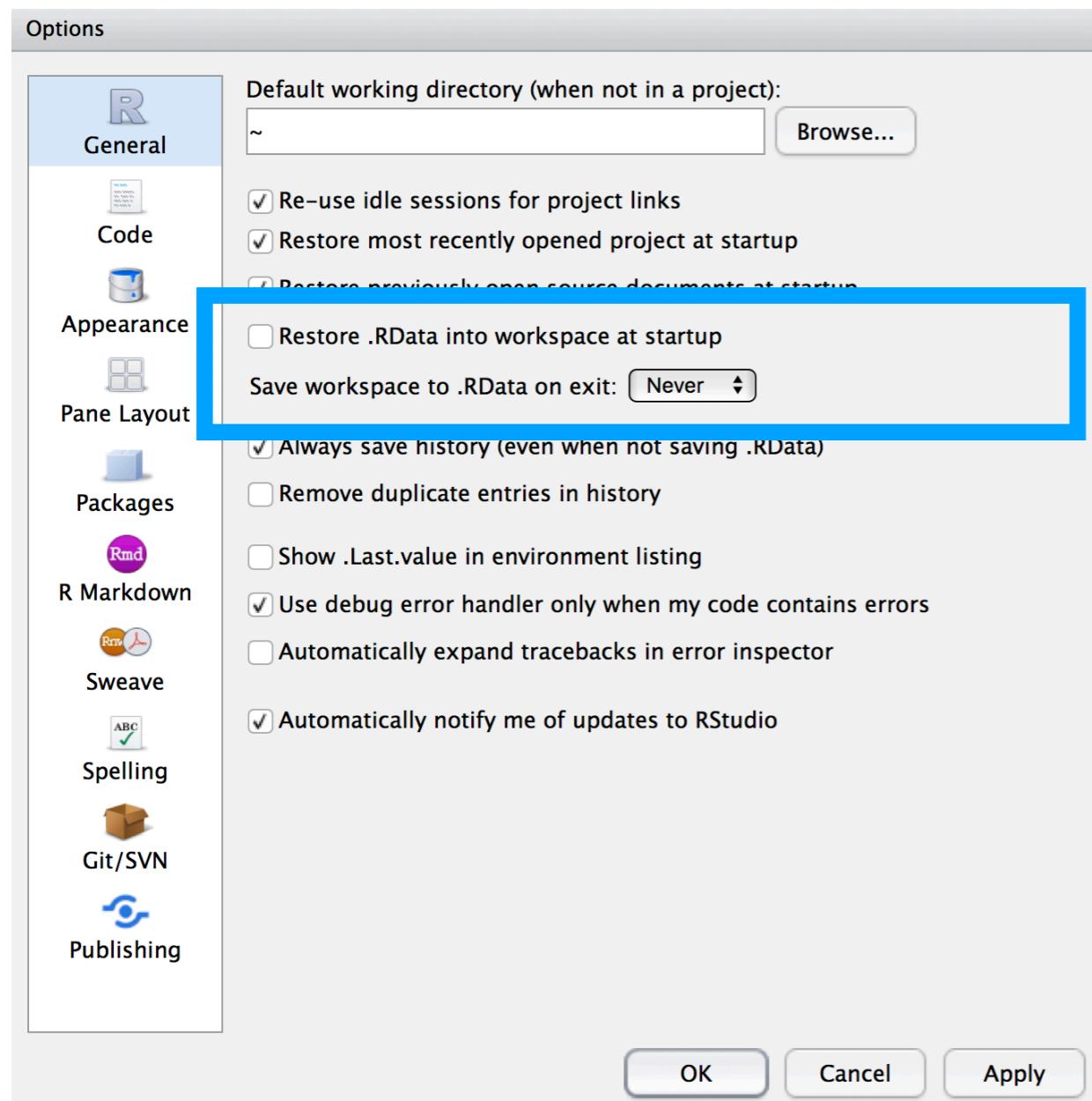
With your R scripts (and your data files), you can recreate the environment. It's much harder to recreate your R scripts from your environment! (...).

To foster this behaviour, I highly recommend that you instruct RStudio not to preserve your workspace between sessions.



<https://r4ds.had.co.nz/index.html>

OPCIONES GLOBALES



AYUDA



Buscar en Google o escribir una URL



Rstats



AYUDA

ChatGPT ▾

↑ Share ⋮

Could you help me with a chunk code in R ?

Of course! Please share the chunk of R code (or describe what you're trying to do), and let me know what issues you're running into. Do you want me to debug it, optimize it, or explain how it works?



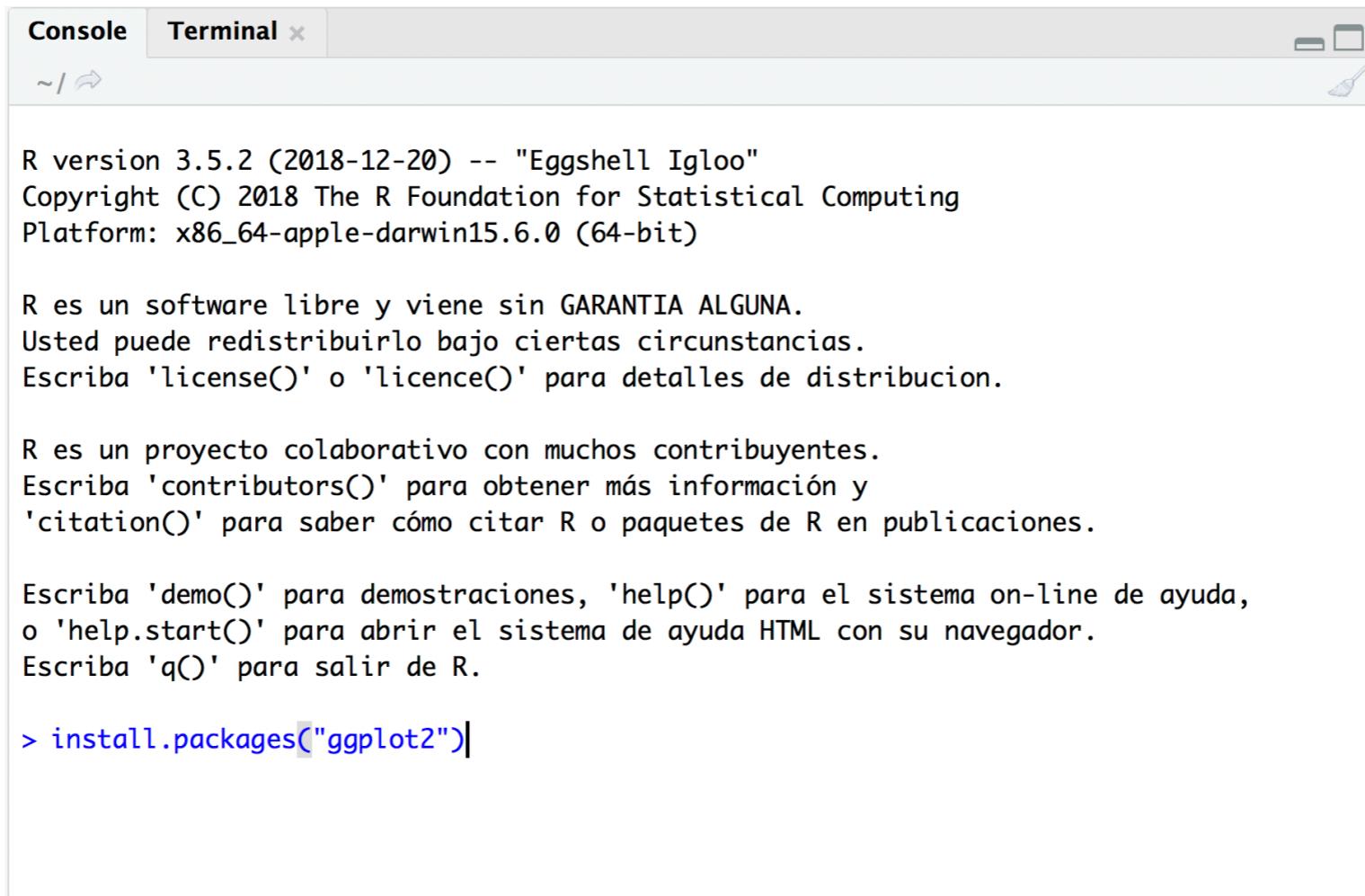
Is this conversation helpful so far?

INSTALACIÓN DE PAQUETES R

- **install.packages()** , función para instalar paquetes en R

INSTALACIÓN DE PAQUETES R

- **install.packages()** , función para instalar paquetes en R



The screenshot shows an R console window with two tabs: "Console" and "Terminal". The "Console" tab is active, displaying the R startup message:

```
R version 3.5.2 (2018-12-20) -- "Eggshell Igloo"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

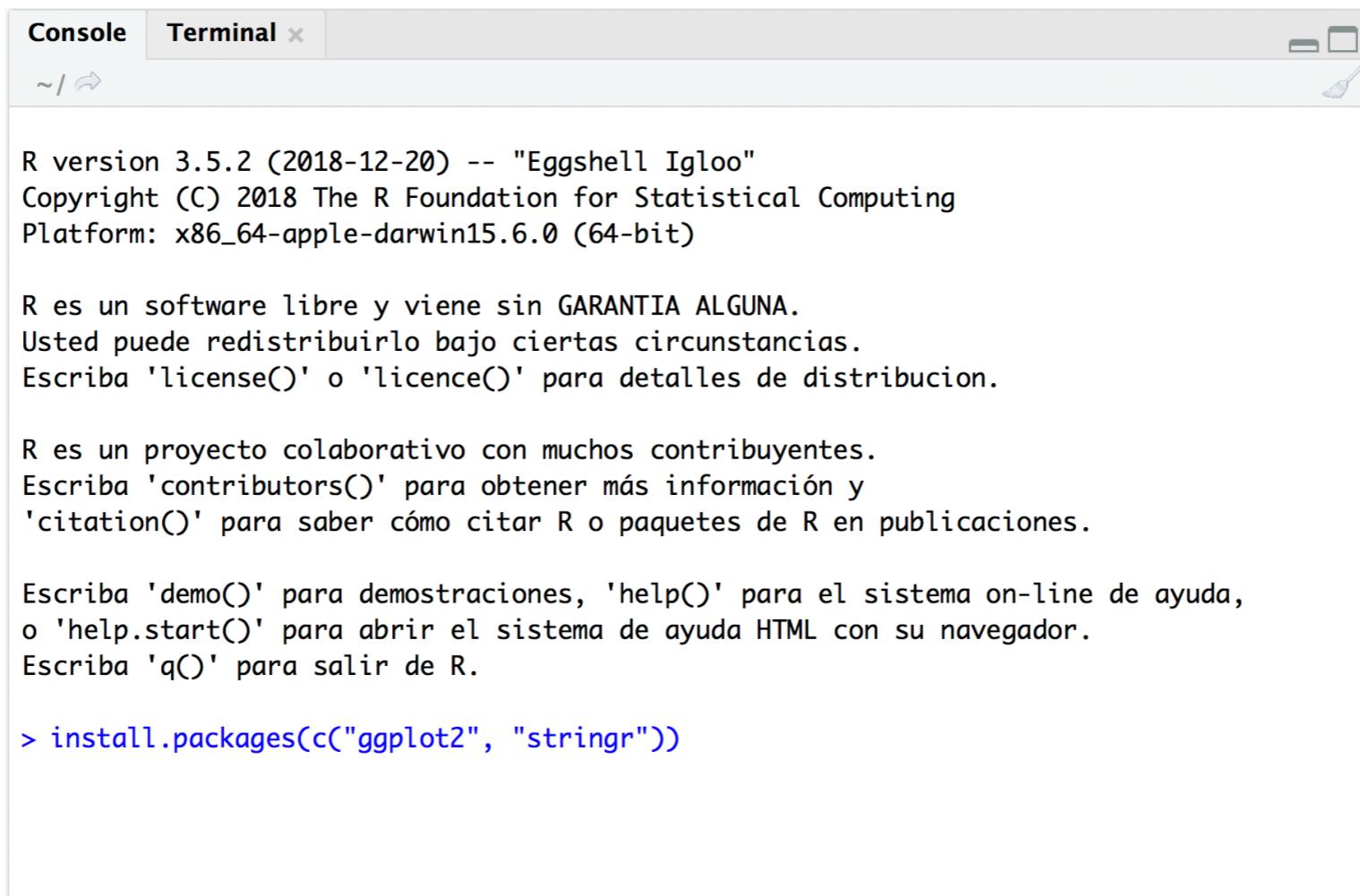
Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.
```

At the bottom of the console, a blue cursor highlights the command:

```
> install.packages("ggplot2")
```

INSTALACIÓN DE PAQUETES R

- **install.packages()** , función para instalar paquetes en R



The screenshot shows an R console window with two tabs: "Console" and "Terminal". The "Console" tab is active, displaying the R startup message:

```
R version 3.5.2 (2018-12-20) -- "Eggshell Igloo"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.
```

At the bottom of the console, a blue-highlighted command is shown:

```
> install.packages(c("ggplot2", "stringr"))
```

- **c("elemento1", "elemento2", ... "elementoN")** , vector carácter

INSTALACIÓN DE PAQUETES R

- **installed.packages()** , función para mostrar todos los paquetes que hemos instalado

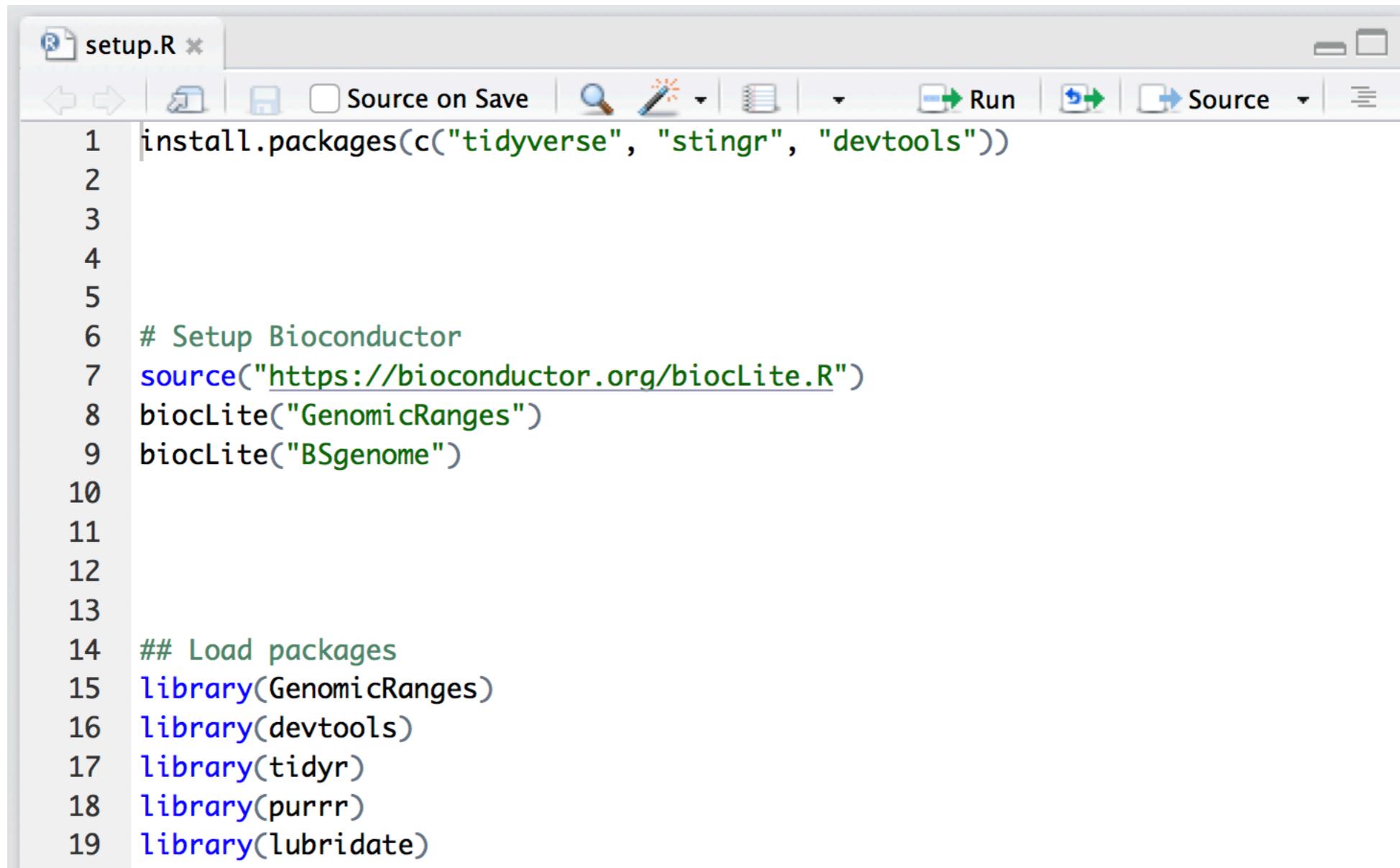
```
Console ~ / 
ESCRIDA license() o licence() para detalles de distribución.

R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.

> installed.packages()
      Package
acepack           "acepack"
ade4              "ade4"
affy              "affy"
affyio             "affyio"
annotate          "annotate"
AnnotationDbi     "AnnotationDbi"
AnnotationHub     "AnnotationHub"
ape                "ape"
assertthat        "assertthat"
backports          "backports"
base               "base"
base64enc          "base64enc"
BH                "BH"
bindr              "bindr"
bindrcpp           "bindrcpp"
Biobase            "Biobase"
BiocGenerics       "BiocGenerics"
BiocInstaller      "BiocInstaller"
BiocParallel        "BiocParallel"
biomaRt            "biomaRt"
Biostrings          "Biostrings"
biovizBase          "biovizBase"
bit                 "bit"
bit64              "bit64"
bitops             "bitops"
blob               "blob"
```

INSTALACIÓN DE PAQUETES R

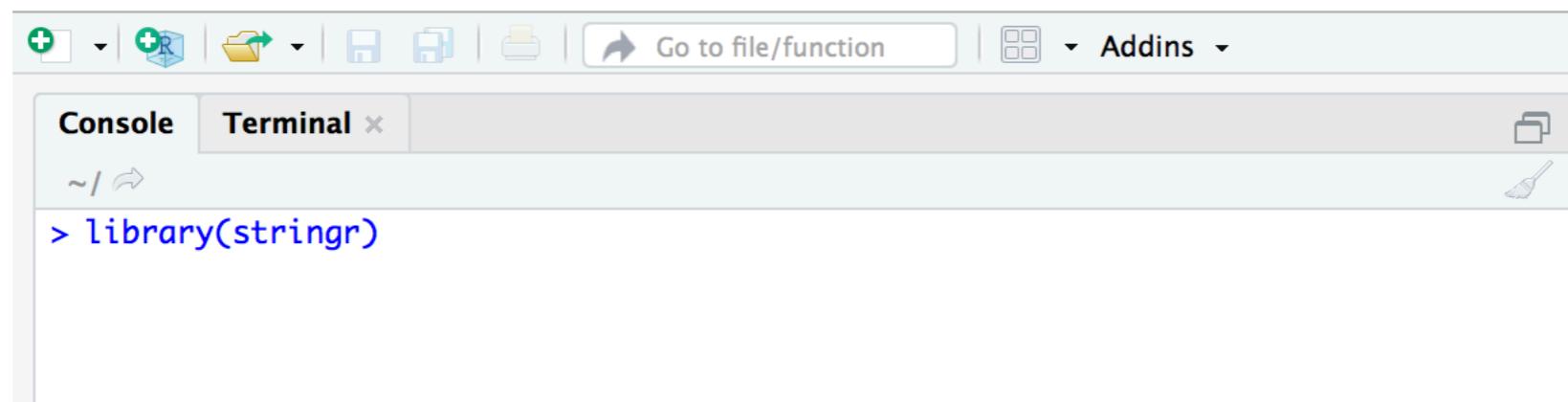


The screenshot shows the RStudio interface with a script editor window titled "setup.R". The window contains R code for package installation and Bioconductor setup. The code is numbered from 1 to 19. The "Run" button in the toolbar is highlighted.

```
1 install.packages(c("tidyverse", "stringr", "devtools"))
2
3
4
5
6 # Setup Bioconductor
7 source("https://bioconductor.org/biocLite.R")
8 biocLite("GenomicRanges")
9 biocLite("BSgenome")
10
11
12
13
14 ## Load packages
15 library(GenomicRanges)
16 library(devtools)
17 library(tidyr)
18 library(purrr)
19 library(lubridate)
```

INSTALACIÓN DE PAQUETES R

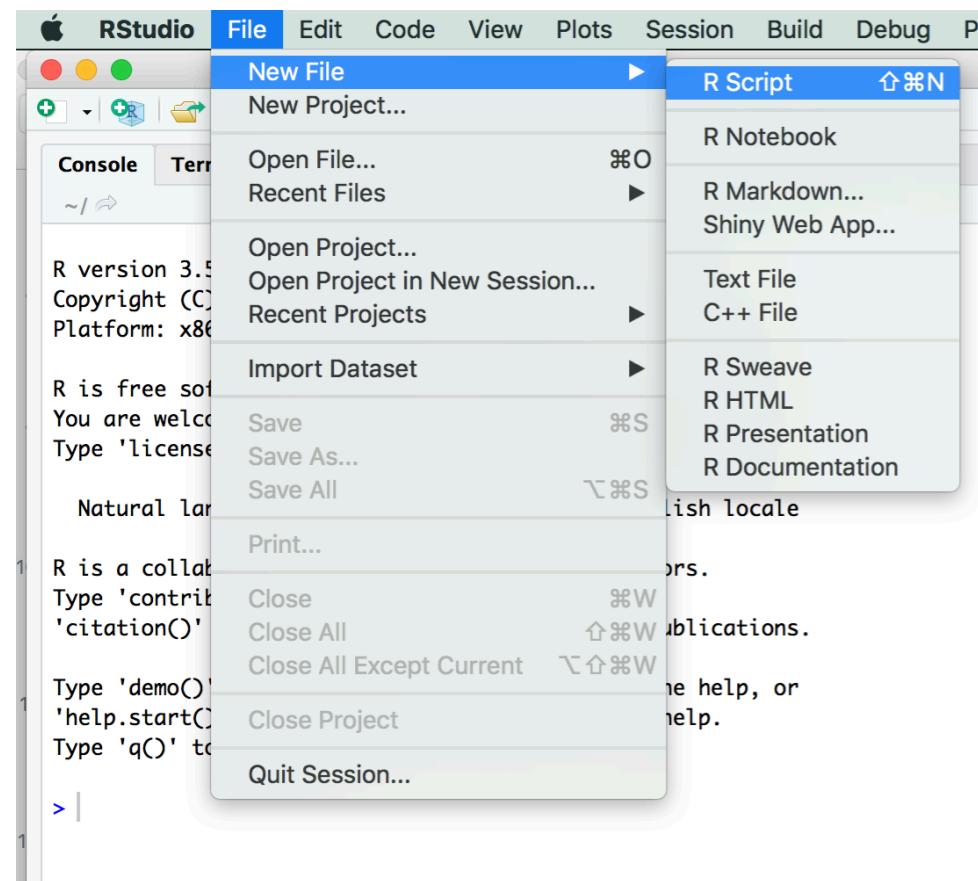
- **library()**, función para cargar en R los paquetes instalados



A screenshot of the RStudio 'Packages' tab. The tab bar at the top has 'Files', 'Plots', 'Packages', 'Help', and 'Viewer' tabs, with 'Packages' highlighted. Below the tab bar is a search bar containing the text 'str'. A red arrow points from the 'library(stringr)' command in the previous screenshot to this one. The main area displays a table of packages:

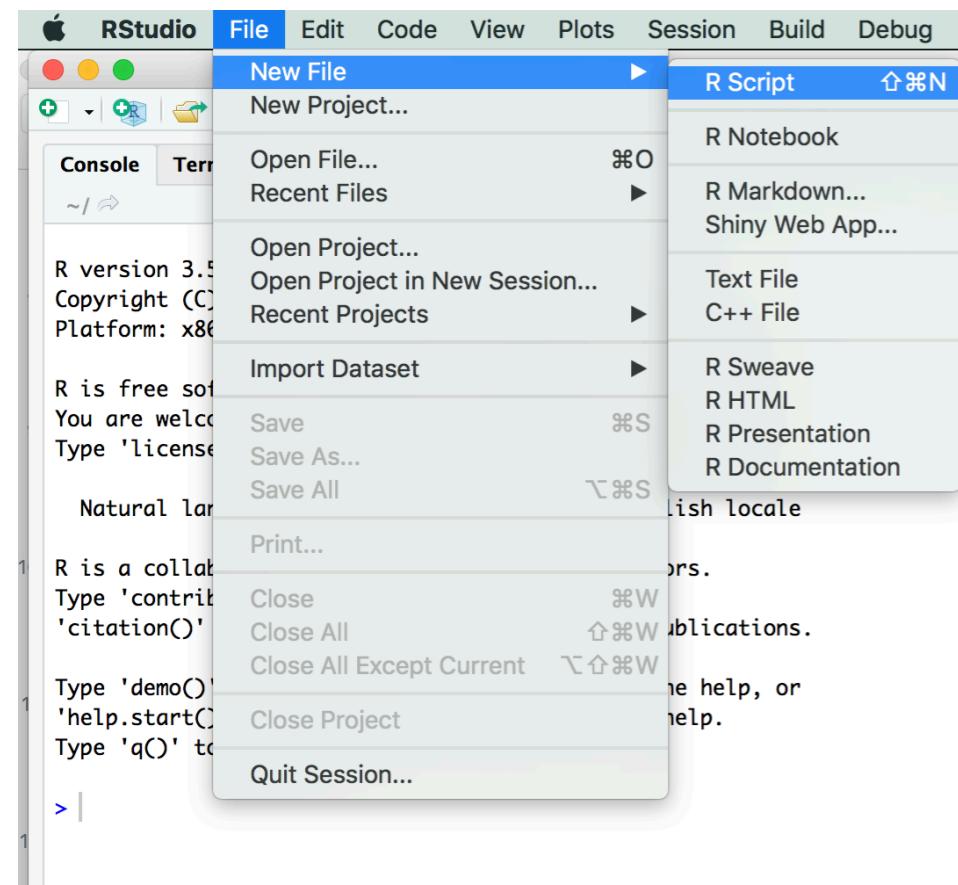
Name	Description	Version
stringi	Character String Processing Facilities	1.2.3
<input checked="" type="checkbox"/> stringr	Simple, Consistent Wrappers for Common String Operations	1.3.1
boot	Bootstrap Functions (Originally by Angelo Canty for S)	1.3-20
glue	Interpreted String Literals	1.2.0
IRanges	Infrastructure for manipulating intervals on sequences	2.14.10
promises	Abstractions for Promise-Based Asynchronous Programming	1.0.1
tidyselect	Select from a Set of Strings	0.2.4

USANDO RSTUDIO POR PRIMERA VEZ



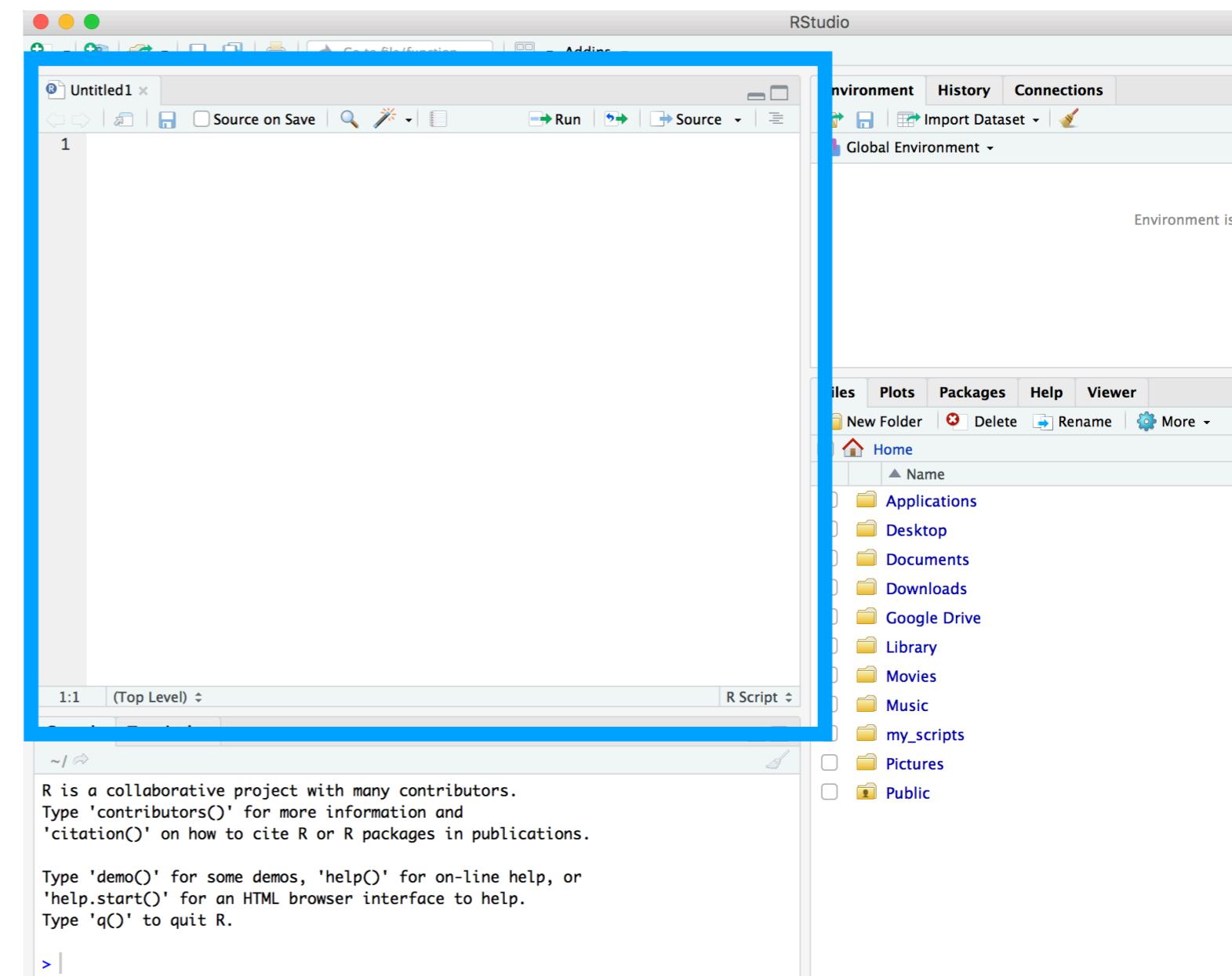
FUNDAMENTOS BÁSICOS DE R

USANDO RSTUDIO POR PRIMERA VEZ



Windows

control-shift-N



Mac

commando-shift-N

CORRIENDO SCRIPTS

The screenshot shows the RStudio IDE interface. On the left, the script editor displays the code for 'mi_primer_script.R'. The code loads the 'datasets' library and plots chicken weight against diet. On the right, the environment browser shows the global environment. Below it, the file browser lists various folders like Applications, Desktop, and my_scripts. The bottom section contains the console and terminal panes.

```
1 # Load datasets library
2 library(datasets)
3
4 # Read data set
5 mydata = ChickWeight
6
7 # Plot weight over diets
8 boxplot(mydata$weight ~ mydata$Diet,
9         xlab = "Diet", ylab = "chickpen weight (g)")
10
11 # Overlay points
12 colors = as.numeric(ChickWeight$Diet)
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
14         col= alpha(colors, 0.4),
15         pch=19)
16
17
```

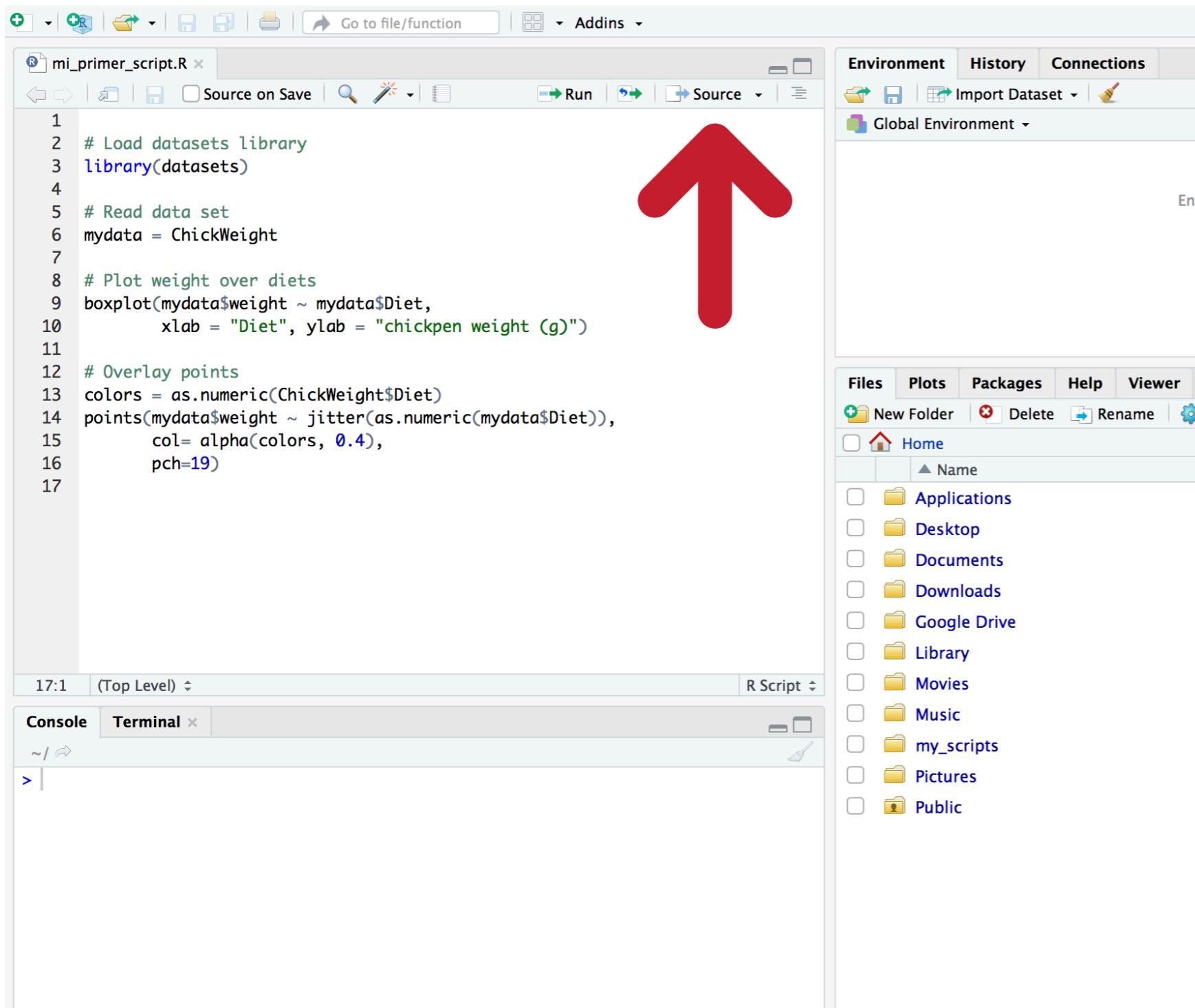
17:1 (Top Level) R Script

Console Terminal

~/

>

CORRIENDO SCRIPTS



```
1 # Load datasets library  
2 library(datasets)  
3  
4 # Read data set  
5 mydata = ChickWeight  
6  
7 # Plot weight over diets  
8 boxplot(mydata$weight ~ mydata$Diet,  
9          xlab = "Diet", ylab = "chickpen weight (g)")  
10  
11 # Overlay points  
12 colors = as.numeric(ChickWeight$Diet)  
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),  
14         col= alpha(colors, 0.4),  
15         pch=19)  
16  
17
```

17:1 (Top Level) R Script

Console Terminal

~/>

Windows

control-shift-enter

Mac

commando-shift-return

CORRIENDO SCRIPTS

The screenshot shows the RStudio interface with the following components:

- Script Editor:** Displays the R script `mi_primer_script.R` containing code to load a dataset, read a data set, plot weight over diets, and overlay points.
- Environment View:** Shows the global environment with a dataset named `mydata` containing 578 observations of 4 variables.
- Plots View:** Displays a boxplot of chicken weight (g) versus Diet (1, 2, 3, 4). The y-axis ranges from 50 to 350. The x-axis is labeled "Diet". The plot shows four distinct groups of data points, each with a different color (grey, red, green, blue) corresponding to the diet group.
- Console View:** Shows the R command history, including the execution of the script's commands.

```

1  # Load datasets library
2  library(datasets)
3
4  # Read data set
5  mydata = ChickWeight
6
7  # Plot weight over diets
8  boxplot(mydata$weight ~ mydata$Diet,
9          xlab = "Diet", ylab = "chickpen weight (g)")
10
11 # Overlay points
12 colors = as.numeric(ChickWeight$Diet)
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
14         col= alpha(colors, 0.4),
15         pch=19)
16
17
2:24 (Top Level) ▾ R Script ▾
Console Terminal ▾
~/
> boxplot(mydata$weight ~ mydata$Diet,
+           xlab = "Diet", ylab = "chickpen weight (g)")

> # Overlay points
> colors = as.numeric(ChickWeight$Diet)

> points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
+         col= alpha(colors, 0.4),
+         pch=19)
>

```

CORRIENDO SCRIPTS

The screenshot shows the RStudio interface with the following components:

- Top Bar:** Includes file navigation icons, a "Go to file/function" search bar, and an "Addins" dropdown.
- Project Bar:** Shows "Project: (None)".
- Code Editor:** Displays the script "mi_primer_script.R" with the following code:

```

1 # Load datasets library
2 library(datasets)
3
4 # Read data set
5 mydata = ChickWeight
6
7 # Plot weight over diets
8 boxplot(mydata$weight ~ mydata$Diet,
9         xlab = "Diet", ylab = "chickpen weight (g)")
10
11 # Overlay points
12 colors = as.numeric(ChickWeight$Diet)
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
14         col= alpha(colors, 0.4),
15         pch=19)
16
17

```
- Environment Tab:** Shows the "Global Environment" tab with "mydata" (578 obs. of 4 variables) and "colors" (num [1:578] 1 1 1 1 1 1 1 1 1 ...).
- Plots Tab:** Contains a boxplot titled "chickpen weight (g)" versus "Diet" (1, 2, 3, 4). The plot shows four groups of data points colored by diet: Diet 1 (grey), Diet 2 (red), Diet 3 (green), and Diet 4 (blue). Individual data points are overlaid on the boxplots using the jittered color scheme defined in the script.
- Console Tab:** Shows the R command history:

```

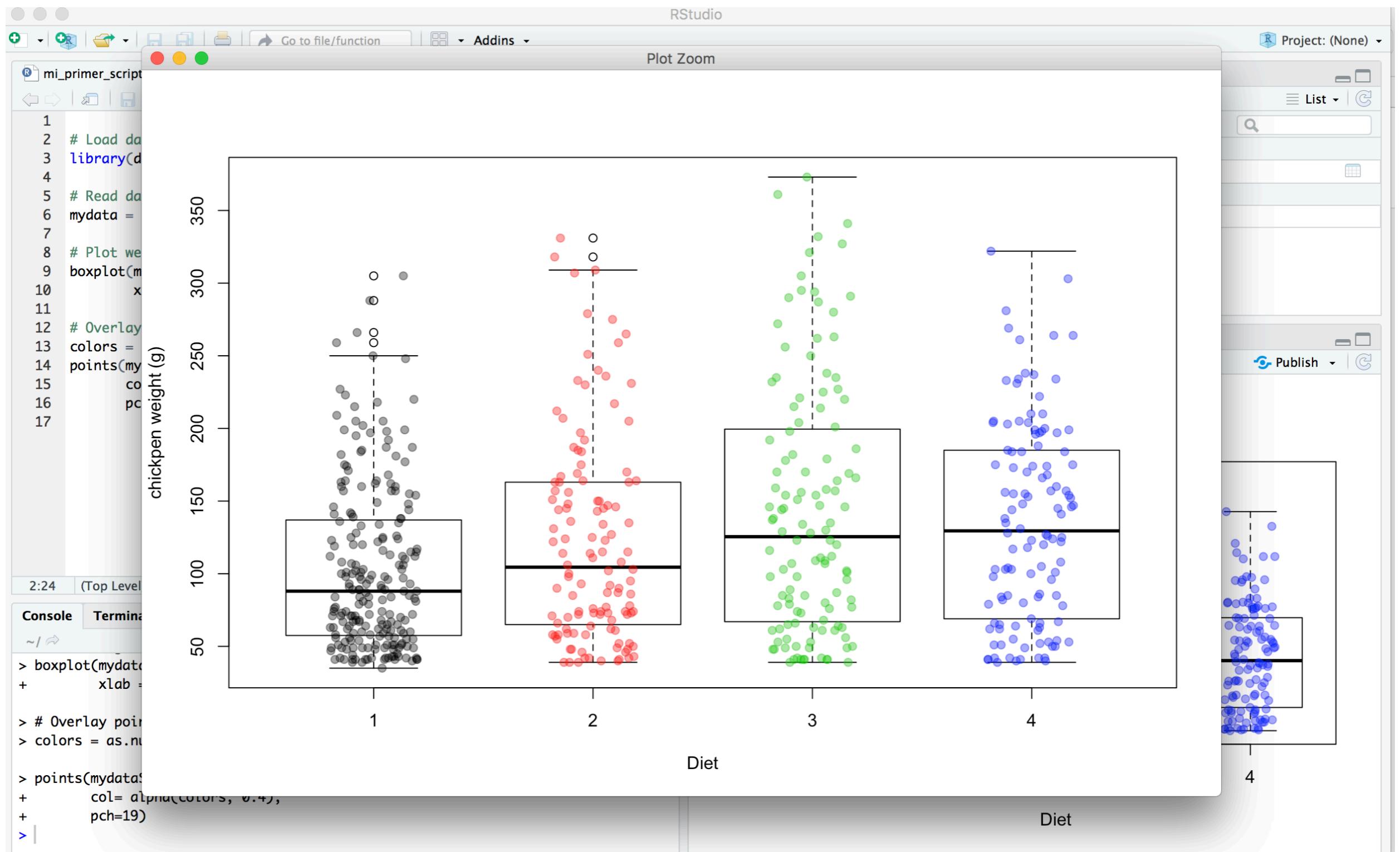
> boxplot(mydata$weight ~ mydata$Diet,
+           xlab = "Diet", ylab = "chickpen weight (g)")

> # Overlay points
> colors = as.numeric(ChickWeight$Diet)

> points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
+         col= alpha(colors, 0.4),
+         pch=19)
>

```

CORRIENDO SCRIPTS



CORRIENDO SCRIPTS

Screenshot of the RStudio IDE interface showing a script file, a plot window, and a preview window.

Script Editor:

```

1 # Load datasets library
2 library(datasets)
3
4 # Read data set
5 mydata = ChickWeight
6
7 # Plot weight over diets
8 boxplot(mydata$weight ~ mydata$Diet,
9         xlab = "Diet", ylab = "chickpen weight (g)")
10
11 # Overlay points
12 colors = as.numeric(ChickWeight$Diet)
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
14        col= alpha(colors, 0.4),
15        pch=19)
16
17

```

Console:

```

> boxplot(mydata$weight ~ mydata$Diet,
+           xlab = "Diet", ylab = "chickpen weight (g)")

> # Overlay points
> colors = as.numeric(ChickWeight$Diet)

> points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
+           col= alpha(colors, 0.4),
+           pch=19)
>

```

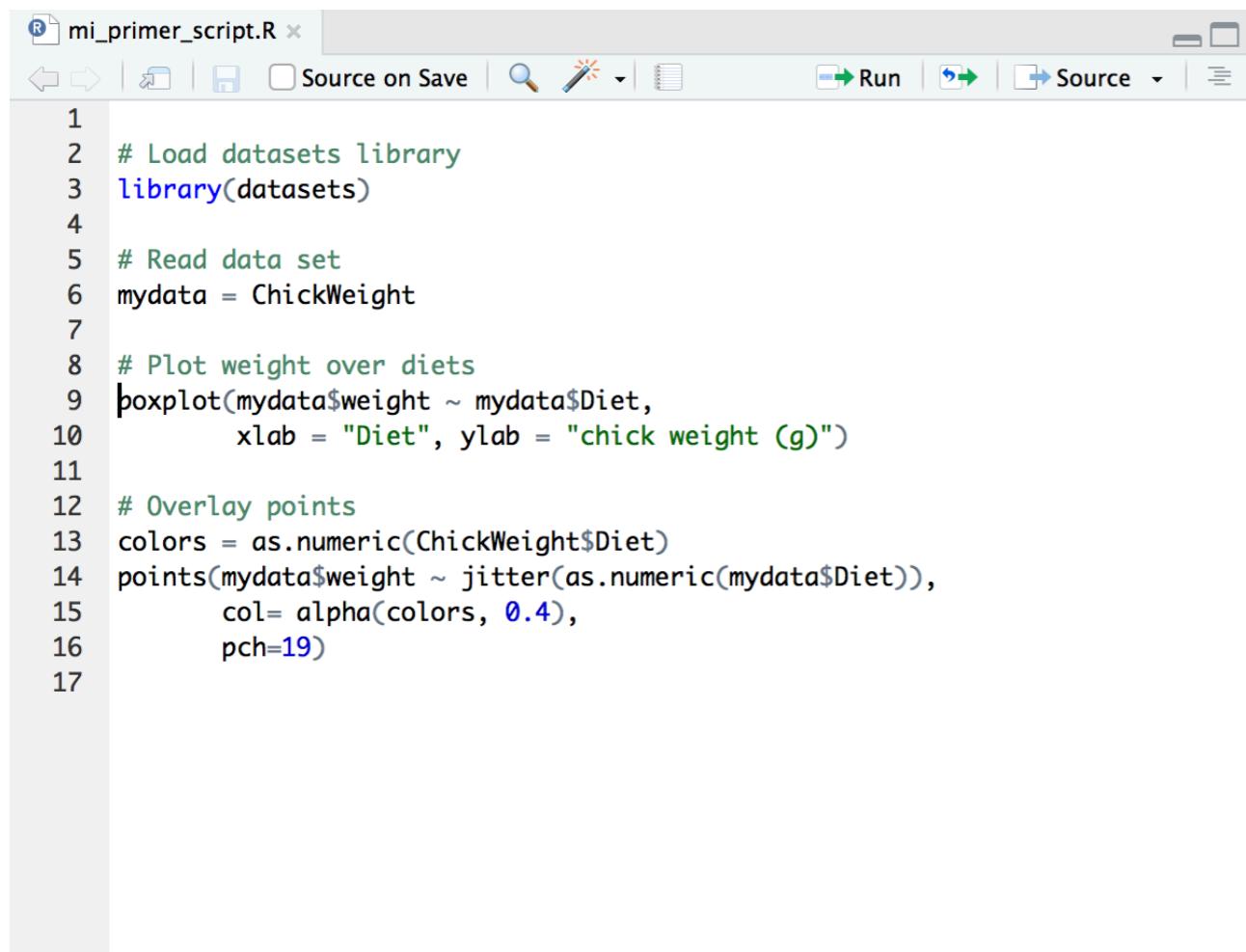
Plot Window (Save Plot as Image dialog):

The dialog shows a boxplot of chickpen weight (g) vs Diet (1, 2, 3, 4). The y-axis ranges from 50 to 350. The x-axis is labeled "Diet". Points are overlaid on the boxplots, colored by diet group (black, red, green, blue).

Preview Window:

A second window displays the same boxplot and overlaid points for Diet groups 3 and 4, suggesting a zoomed-in view or a duplicate plot.

CORRIENDO SCRIPTS



```
1 # Load datasets library
2 library(datasets)
3
4 # Read data set
5 mydata = ChickWeight
6
7 # Plot weight over diets
8 boxplot(mydata$weight ~ mydata$Diet,
9         xlab = "Diet", ylab = "chick weight (g)")
10
11 # Overlay points
12 colors = as.numeric(ChickWeight$Diet)
13 points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
14        col= alpha(colors, 0.4),
15        pch=19)
16
17
```

Windows

control-enter



```
9:1 (Top Level) ▾ R Script ▾
Console Terminal ×
~/
+     xlab = "Diet", ylab = "chick weight (g)")
> # Overlay points
> colors = as.numeric(ChickWeight$Diet)
> points(mydata$weight ~ jitter(as.numeric(mydata$Diet)),
+        col= alpha(colors, 0.4),
+        pch=19)
> # Plot weight over diets
> boxplot(mydata$weight ~ mydata$Diet,
+         xlab = "Diet", ylab = "chick weight (g)")
>
```

Mac

comando-return

RESUMEN

- **install.packages()** , función para instalar paquetes en R
- **library()** , función para cargar en R los paquetes instalados
- **installed.packages()** , función para mostrar todos los paquetes que hemos instalado