

R PARA PROCESAMIENTO DE DATOS

1.0 - ENTORNO R Y RSTUDIO

Moisés Gualapuro, Cristina Ramos

2020-09-28



Moisés Gualapuro - `moises.gualapuro@ikiam.edu.ec` - Universidad Regional Amazónica IKIAM, Tena (Ecuador)

Cristina Ramos - `cristina.ramos@ikiam.edu.ec` - Universidad Regional Amazónica IKIAM, Tena (Ecuador)

CONTENIDO

1. RESULTADOS DE APRENDIZAJE

Al finalizar la unidad, el estudiante será capaz de:

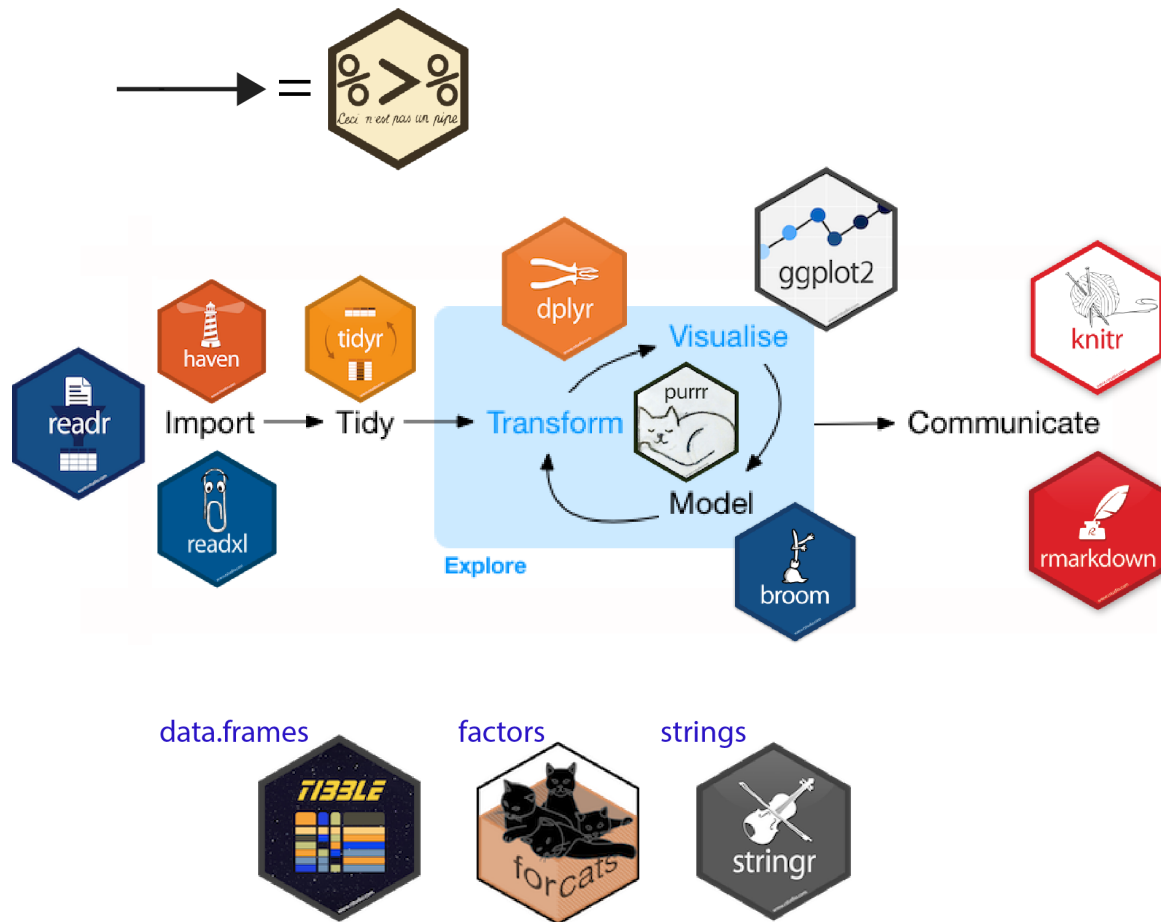
- Instalar R y Rstudio en el equipo personal
- Crear su usuario en Rstudio cloud (<https://rstudio.cloud/>)
- Tener familiaridad con el entorno de Rstudio

2. ANTECEDENTES

R es un lenguaje de programación para estadística computacional. Es un paquete de software libre disponible bajo GNU Licencia General Pública. R y sus librerías están diseñadas para estadística y visualización, entre ellas las estadísticas descriptivas, inferencias estadísticas y análisis de regresión.

2.1 Ventajas del uso de R

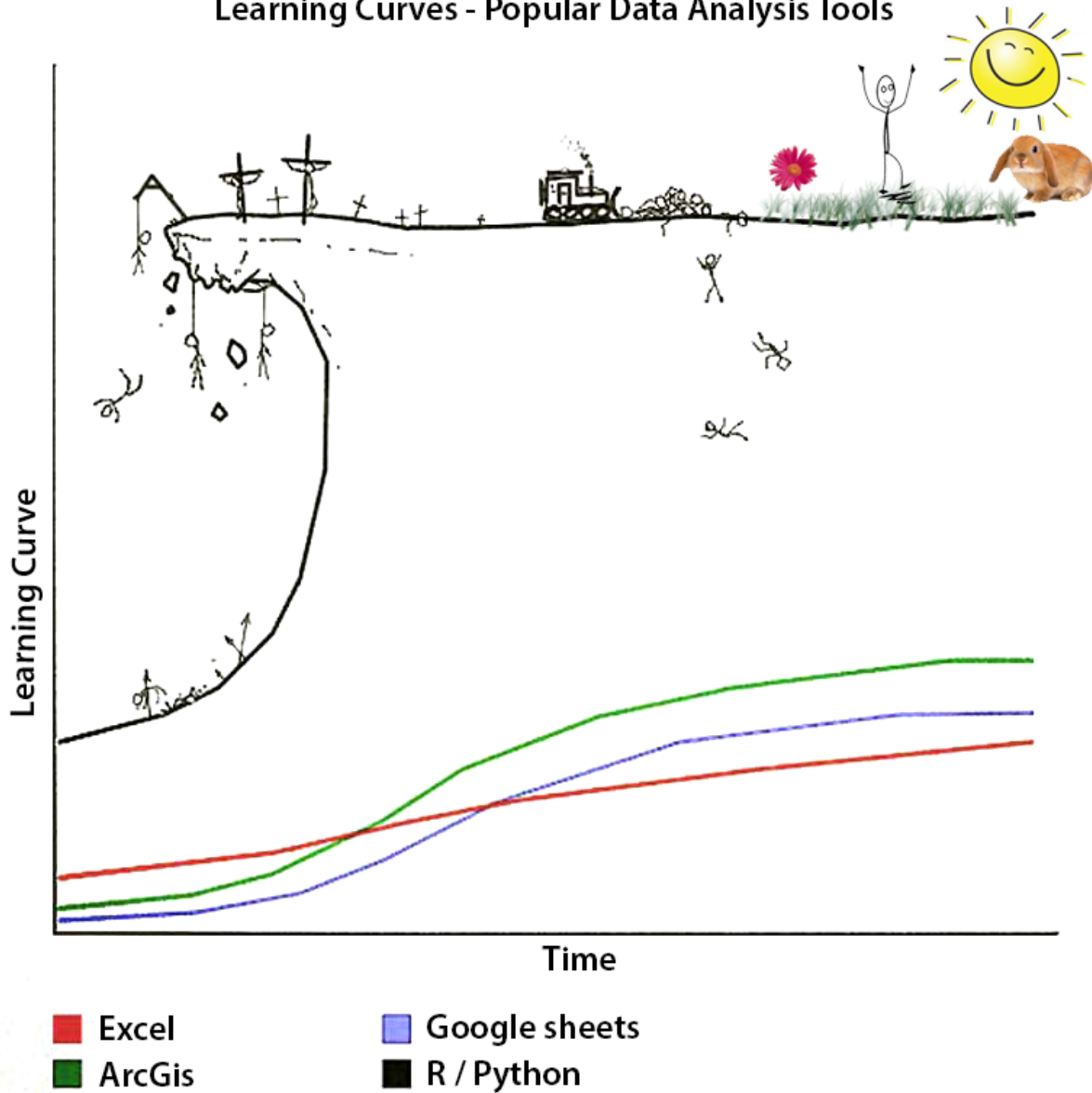
- Es compatible con todos los sistemas operativos (Linux, MacOS, Windows).
- Es un entorno integrado, especialmente desarrollado para el análisis de datos, los cálculos estadísticos y las representaciones gráficas.
- Es un lenguaje de programación sencillo.
- Es **software libre**.
- Ampliamente usado en la investigación científica.
- Permite crear visualizaciones de datos de alta calidad (ggplot).
- Permite crear dashboards para visualizar y analizar datos (Markdown, rKnit, Latex).
- Permite crear informes automáticos.
- Dispone de herramientas de análisis estadístico para ahondar en el conocimiento de los datos.
- Es reproducible.
- Interactividad en la visualización (shiny: <https://shiny.rstudio.com/gallery/>, plot.ly: <https://plot.ly/r/line-and-scatter/>).
- Comunidad de apoyo (StackOverflow: <https://stackoverflow.com/questions/tagged/r>, cross Validated: <https://stats.stackexchange.com/>, R Help Mailing List: <https://stat.ethz.ch/mailman/listinfo/r-help>, r-bloggers: <https://www.r-bloggers.com/>).
- Amplia disponibilidad de paquetes (https://cran.r-project.org/web/packages/available_packages_by_name.html, **tidyverse**: <https://www.tidyverse.org/>).
- Amplia disponibilidad de paquetes para ciencias de la vida (<https://www.bioconductor.org/>).



2.2 Desventajas del uso de R

- R podría consumir toda la memoria RAM disponible, dificulta el trabajo con datos masivos.
- No se puede reclamar si algo no funciona.
- La calidad de algunos paquetes puede no ser ideal.
- Al inicio no es user-frienly. La curva de aprendizaje es lenat al inicio, luego es exponencial.
- Fácil equivocarse y no reconocer dónde se encuentra el error.

Learning Curves - Popular Data Analysis Tools



3. INSTALAR R Y RSTUDIO

3.1 Instalar R

R se puede descargar de <https://cran.r-project.org>; luego se debe seguir las instrucciones e instalar con las opciones predeterminadas.



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

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

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

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

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.

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 (2020-06-22, Taking Off Again) [R-4.0.2.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension [packages](#)

Questions About R

- If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

What are R and CRAN?

En Ubuntu se ejecuta en el terminal: `sudo apt-get install r-base`.

```
moisesgualapuro@moisespc: ~  
(base) moisesgualapuro@moisespc: $ sudo apt-get install r-base  
[sudo] contraseña para moisesgualapuro:  
Leyendo lista de paquetes... Hecho  
Creando árbol de dependencias  
Leyendo la información de estado... Hecho  
r-base ya está en su versión más reciente (3.6.3-2).  
Los paquetes indicados a continuación se instalaron de forma automática y ya no son  
necesarios.
```

3.2 Instalar RStudio

Luego de instalar R, se debe instalar RStudio desde: <https://rstudio.com/products/rstudio/download/>; luego se debe seguir las instrucciones e instalar con las opciones predeterminadas.

RStudio Desktop 1.3.1093 - [Release Notes](#)

1. Install R. RStudio requires [R 3.0.1+](#).
2. Download RStudio Desktop. Find your operating system in the table below.



All Installers

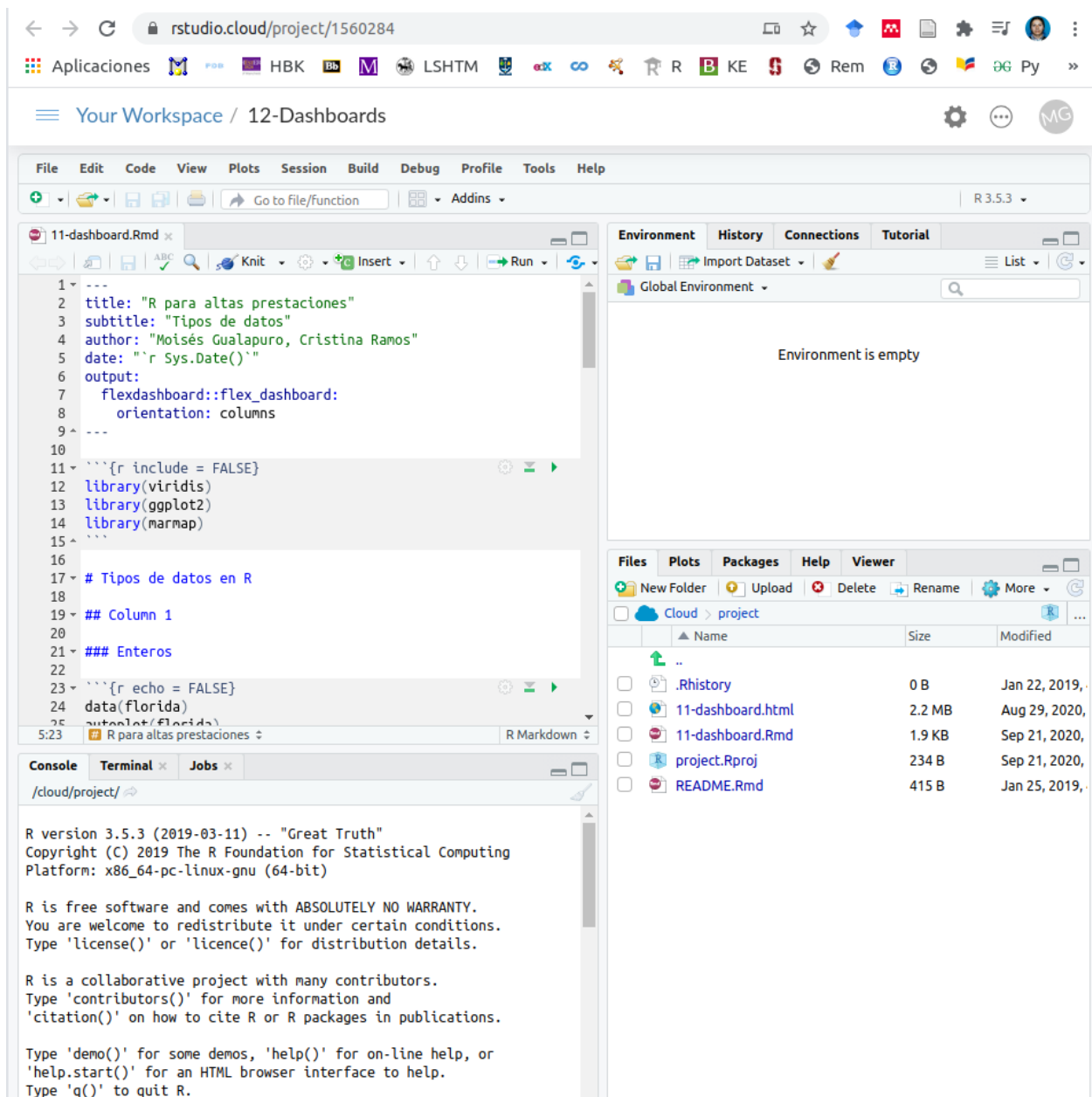
Linux users may need to [import RStudio's public code-signing key](#) prior to installation, depending on the operating system's security policy

RStudio requires a 64-bit operating system. If you are on a 32 bit system, you can use an [older version of RStudio](#).

OS	Download	Size	SHA-256
Windows 10/8/7	RStudio-1.3.1093.exe	171.62 MB	62b9e60a
macOS 10.13+	RStudio-1.3.1093.dmg	148.66 MB	bdc4d3a4
Ubuntu 16	rstudio-1.3.1093-amd64.deb	124.33 MB	72f05048
Ubuntu 18/Debian 10	rstudio-1.3.1093-amd64.deb	126.80 MB	ff222177
Fedora 19/Red Hat 7	rstudio-1.3.1093-x86_64.rpm	146.96 MB	ed1f6ef8
Fedora 28/Red Hat 8	rstudio-1.3.1093-x86_64.rpm	151.05 MB	01a978f3
Debian 9	rstudio-1.3.1093-amd64.deb	127.00 MB	a747f9f9
SLES/OpenSUSE 12	rstudio-1.3.1093-x86_64.rpm	119.43 MB	5016cbcf
OpenSUSE 15	rstudio-1.3.1093-x86_64.rpm	128.40 MB	cf47e32d

3.2.1 RStudio Cloud

Ingresa a <https://rstudio.cloud> y crea un usuario como en cualquier plataforma de correo o red social.



4. ENTORNO R EN EL TERMINAL Y RSTUDIO

4.1 Consola de R o R en el terminal

La primera opción es abrir la consola de R base:

```
R version 3.6.3 (2020-02-29) -- "Holding the Windsock"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (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 publicac
iones.

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

[Previously saved workspace restored]

> 2+2
[1] 4
> a <- 3
> a+2
[1] 5
> print("Bienvenidos al curso de R para altas prestaciones")
[1] "Bienvenidos al curso de R para altas prestaciones"
> █
```

La segunda opción es abrir el terminal (Ctrl + Alt + t) o la línea de comando (cmd)


```
(base) moisesgualapuro@moisespc:~$ R

R version 3.6.3 (2020-02-29) -- "Holding the Windsock"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (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.

r[Previously saved workspace restored]

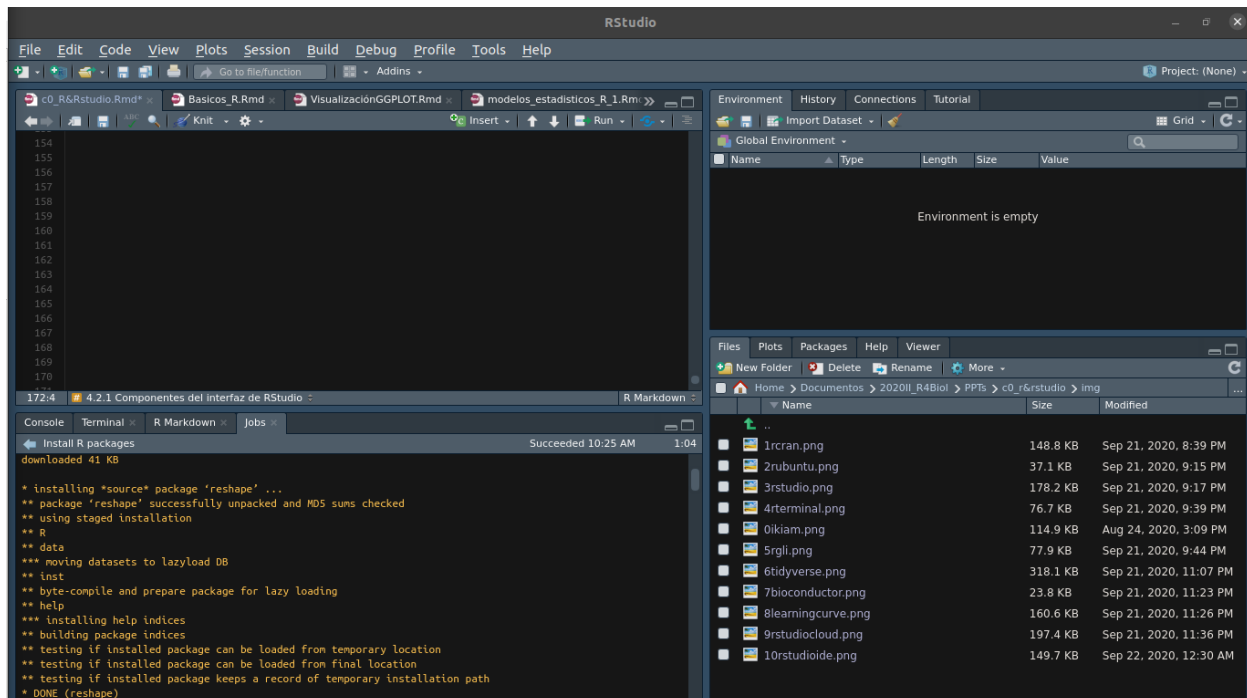
> print("Bienvenido al curso de R para altas prestaciones")
[1] "Bienvenido al curso de R para altas prestaciones"
>
```

4.2 R en RStudio

La tercera opción es usar R con RStudio.

“RStudio es un *integrated development environment* (IDE) para R. Incluye una consola, editor con resaltador sintáctico que soporta la ejecución de código, asimismo herramientas de gráfica, registro histórico, debugging y administración del entorno de trabajo (workspace)”

Al abrir RStudio se tendrá la siguiente ventana:



En el interfaz de RStudio se tiene 4 áreas, además de la barra de opciones en la parte superior.

- **Área Superior - Izquierda:** es el **editor de sintaxis**; se trata del lugar donde editamos la sintaxis para posteriormente ejecutarla. Al escribir allí no sucederá nada, a no ser que se apriete algún botón para ejecutar los comandos o la tecla **ctrl+enter**.
- **Área Superior - Derecha:** es el “**entorno de trabajo**” del programa: en este lugar se muestra el conjunto de datos y los “objetos” (resultados, variables, gráficos, etc.) que se almacenan al ejecutar diferentes análisis.
- **Área Inferior - Derecha:** paneles de **archivos**, de **salida**, de **paquetes** y de **ayuda**;
 - (i) la pestaña **files** permite ver el historial de archivos trabajados con el programa;
 - (ii) la pestaña **plots** permite visualizar los gráficos que se generen;
 - (iii) la pestaña **packages** permite ver los paquetes descargados y guardados en el disco duro así como gestionar su instalación o actualización;
 - (iv) la ventana **help** permite acceder al CRAN - Comprehensive R Archive Network (siempre que se cuente con conexión a Internet), página oficial del software que ofrece diferentes recursos para el programa: manuales para el usuario, cursos on line, información general, descarga de paquetes, información de los paquetes instalados, etc. Esta última pestaña es bastante útil: empleando el motor de búsqueda se accede de manera rápida a manuales de uso de los diferentes paquetes (y sus funciones) instalados en el computador (esto no requiere conexión a Internet).7;
 - (v) la ventana **viewer** muestra los resultados al construir reportes mediante funcionalidades tipo **rmarkdown**.
- **Área Inferior - Izquierda:** es la **consola**. Corresponde a lo que sería el software R en su versión básica. Allí el software ejecuta las operaciones realizadas desde el editor de sintaxis.

4.2.1 Componentes del interfaz de RStudio

Las funcionalidades señaladas a continuación están disponibles en: https://github.com/rstudio/cheatsheets/blob/master/translations/spanish/rstudio-ide_Spanish_Translation_Monica_Alonso.pdf

Documentos y Apps

Abrir archivos Shiny, R Markdown, knitr, Sweave, LaTeX, Rd y más en Panel Fuente

Chequear ortografía, Rendir salida, Elegir formato de salida, Elegir ubicación de salida, Insertar chunk de código

Acceso a guía markdown en: **Help > Markdown Quick Reference**

RStudio reconoce que archivos llamados **app.R**, **server.R**, **ui.R**, y **global.R** corresponden a una **app shiny**

Escribir Código

Navegar en tabulaciones, Abrir en nueva ventana, Salvar, Encontrar y reemplazar, Compilar como notebook, Correr código seleccionado

Cursos de usuarios compartidos, Selección múltiple de cursos / columnas con Alt + mouse drag, Diagnóstico de código que aparece en el margen. Desplácese sobre el símbolo de diagnóstico para más detalles, Sintaxis resaltada basado en extensión del archivo, Tab completa el nombre de función, ruta de archivo, argumentos, y más.

Fragmento de código Multi-lenguaje para usar rápidamente un bloque de código común, Cambiar tipo archivo

Directorio de trabajo, Oprimir para ver historia de comandos, Maximizar/minimizar paneles, Arrastrar bordes de paneles

Soporte R

Importar datos con asistente, Historia de comandos anteriores para correr/copiar, Mostrar presentación diapositivas, **File > New File > R Presentation**

Cargar área de trabajo, Salvar área de trabajo, Eliminar todos los objetos salvados, Buscar dentro del entorno, Mostrar objetos como lista o grilla

Muestra los objetos guardados por tipo con una breve descripción, Ver en visor de datos, Ver código fuente de la función

Crear carpeta, Cargar, Borrar, Renombrar, Cambiar directorio

Ruta al directorio mostrado, un explorador de archivo apuntando al directorio de trabajo. Clicar sobre el archivo o directorio para abrir.

Modo Depuración

Abrir con **debug()**, **browser()**, o un breakpoint. RStudio abrirá el modo depuración cuando encuentre un punto de parada mientras ejecuta el código.

Clicar cerca del número de línea para agregar/sacar un punto de parada.

La línea resaltada indica dónde se pausó la ejecución.

Correr comandos en entorno donde la ejecución está pausada

Examinar variables en entorno de ejecución

Seleccionar función en rastreo para depurar

Pasar por código de a una línea por vez

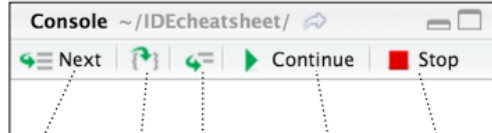
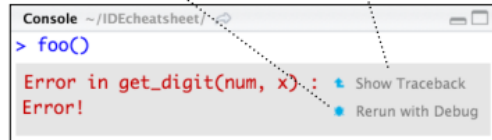
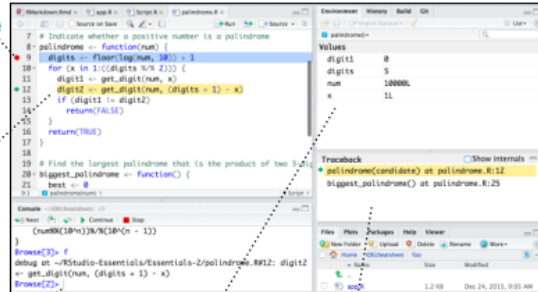
Entrar y salir de las funciones a ejecutar

Reanudar la ejecución

Salir de modo depuración

Entrar a modo depuración desde origen del error

Abrir rastreo para examinar las funciones que R llamó antes que el error ocurriera



5. MAYOR INFORMACIÓN

- Items de cada componente de la interfaz de RStudio: <http://wpd.ugr.es/~bioestad/guia-r-studio/practica-1-r-studio/>
- Textos de R: <https://rstudio.com/resources/books/>
- Guías rápidas (Cheatsheets) de algunos paquetes de R, disponibles en: <https://github.com/rstudio/cheatsheets>