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:

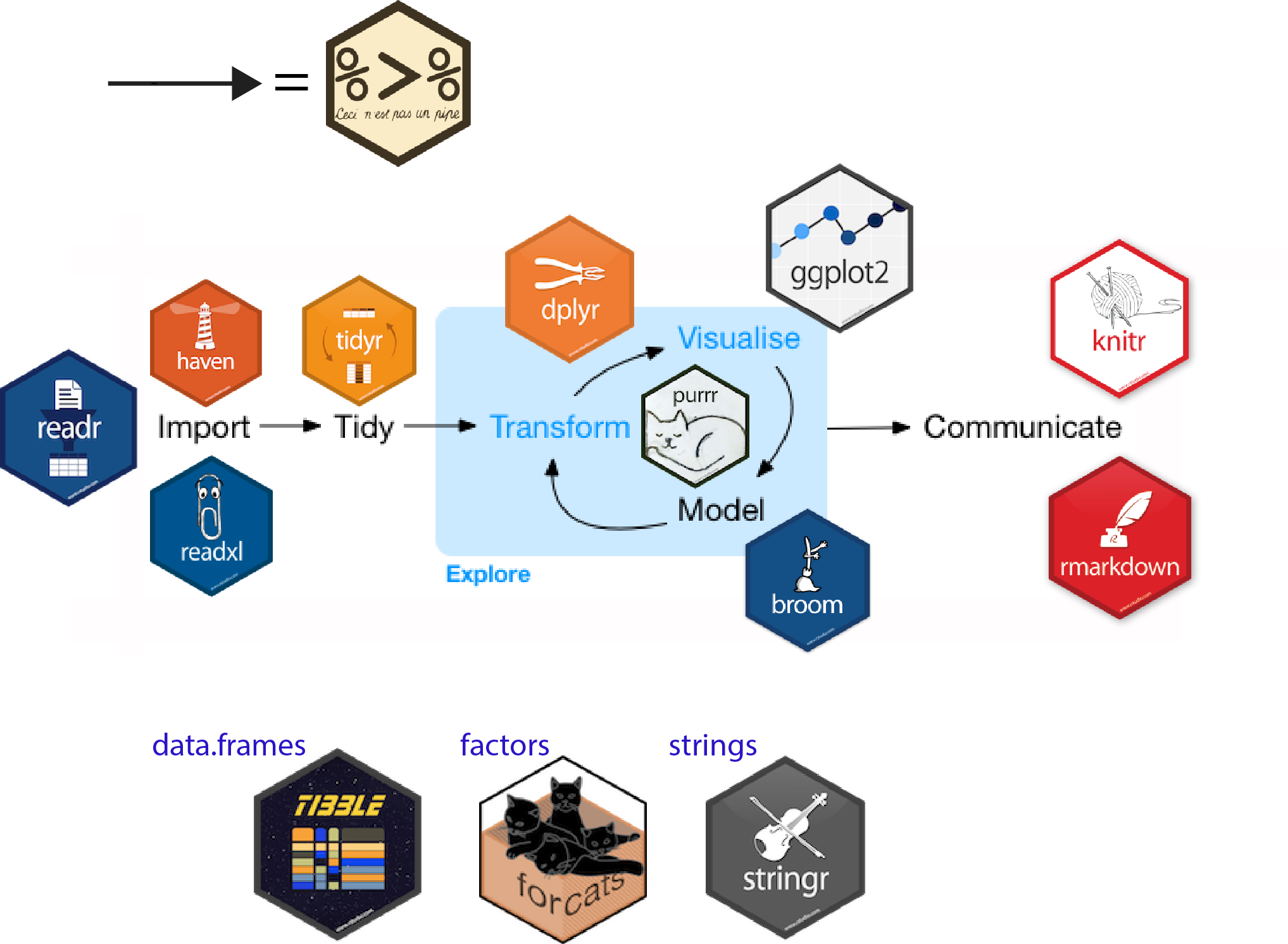
* Instalar R y Rstudio en el equipo personal
* Crear su usuario en Rstudio cloud (<https://rstudio.cloud/>)
* Tener familiaridad con el entorno de Rstudio

# 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.

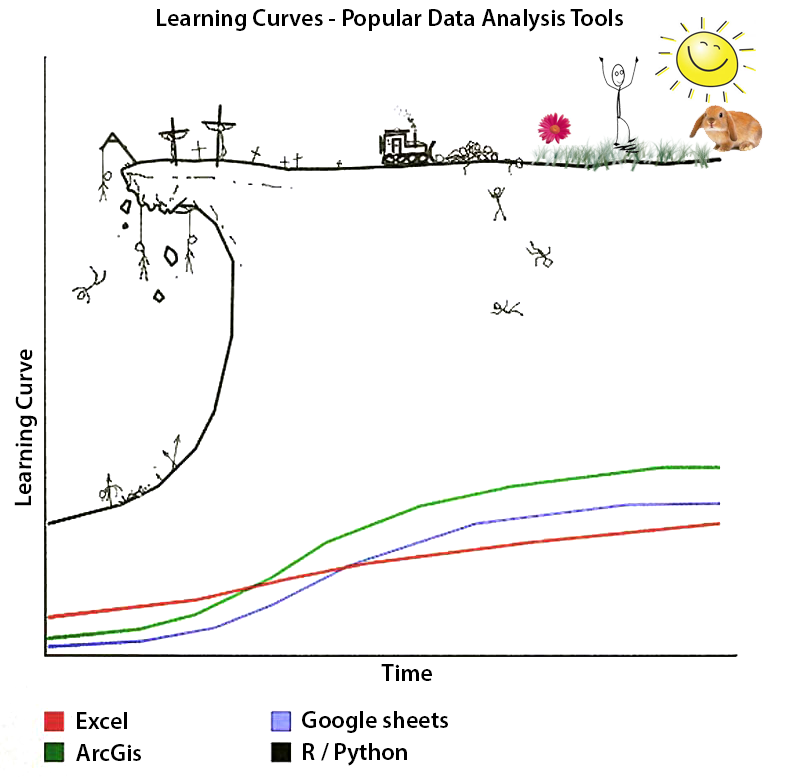
## 2.1 Ventajas del uso de R

* 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.



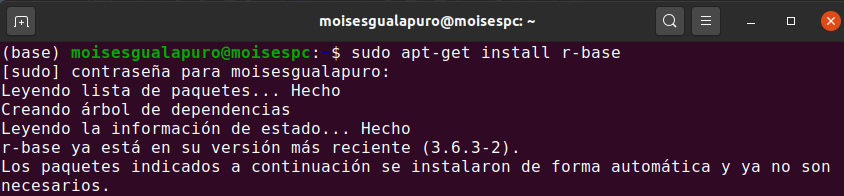
# 3. INSTALAR R Y RSTUDIO

## 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.

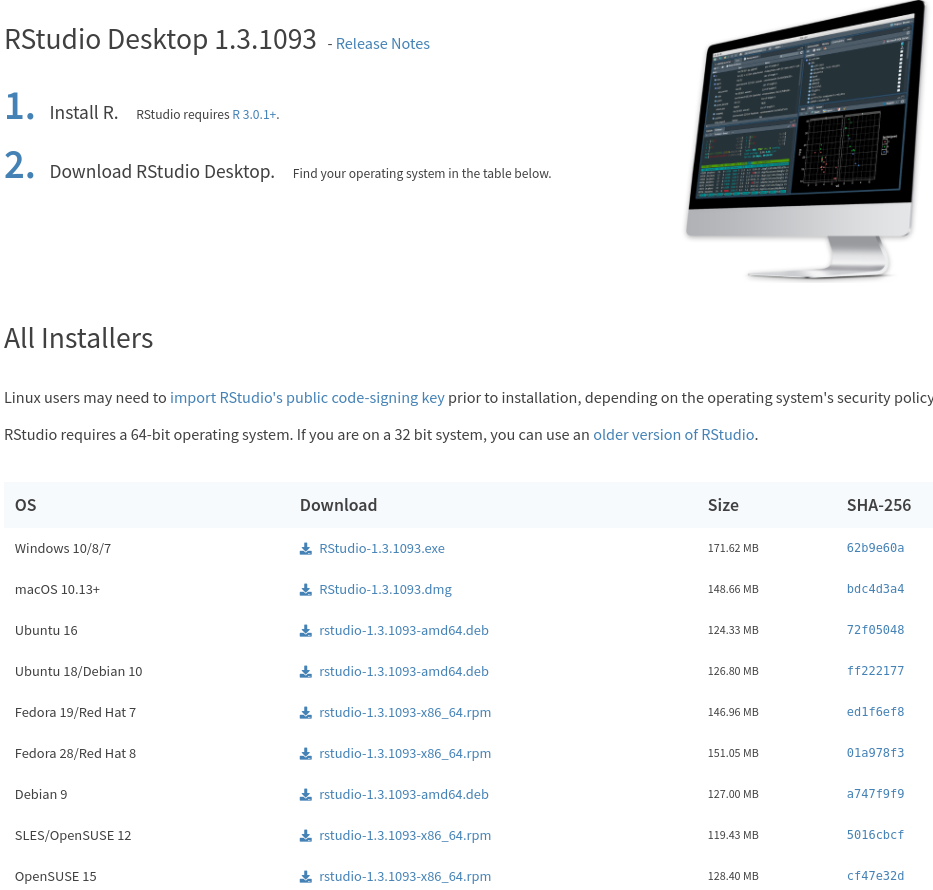


En Ubuntu se ejecuta en el terminal: sudo apt-get install r-base.



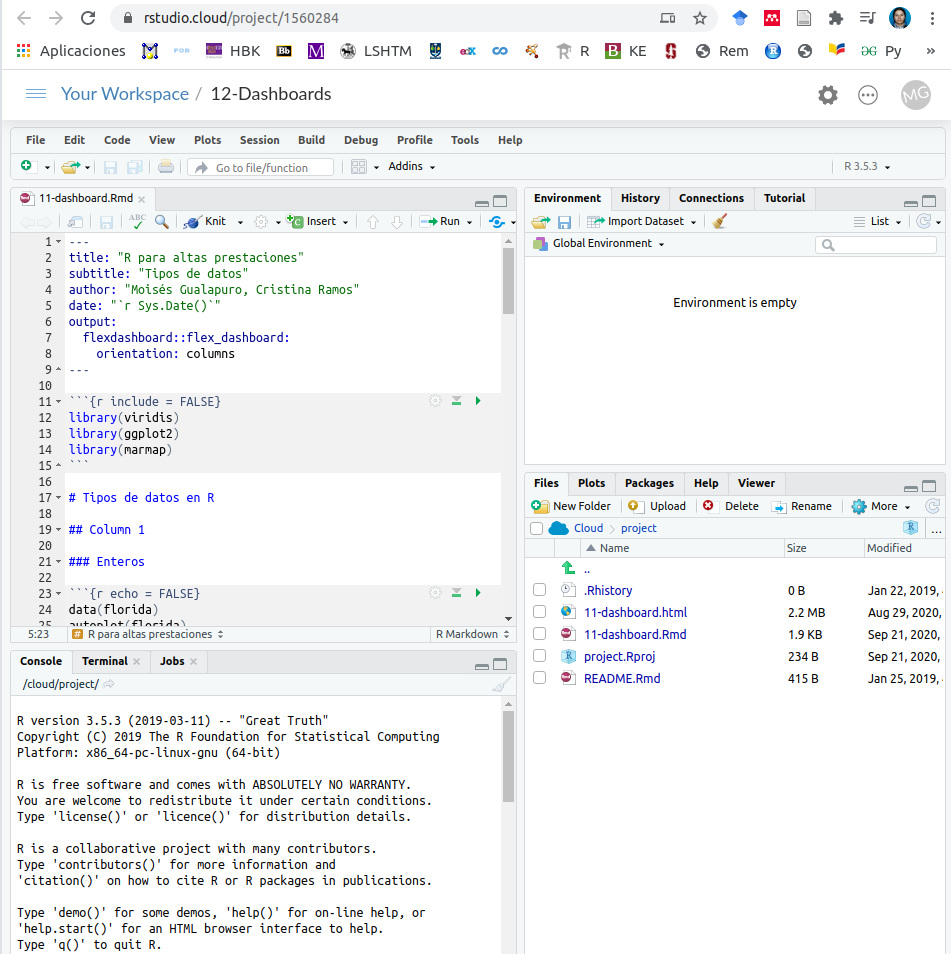
## 3.2 Instalar RStudio

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.



### 3.2.1 RStudio Cloud

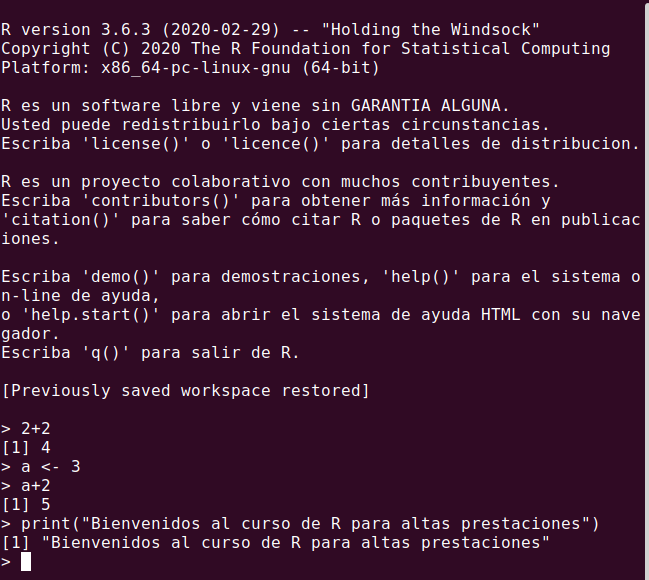
Ingresar a <https://rstudio.cloud> y crear un usuario como en cualquier plataforma de correo o red social.



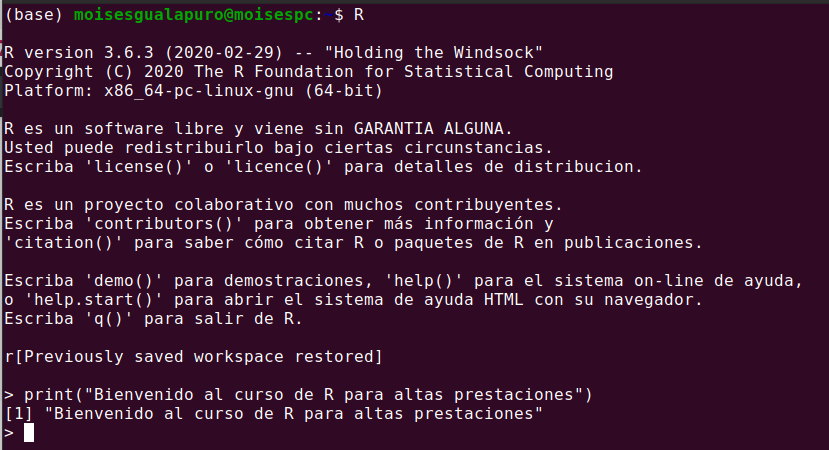
# 4. ENTORNO R EN EL TERMINAL Y RSTUDIO

## 4.1 Consola de R o R en el terminal

La primera opción es abrir la consola de R base:



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

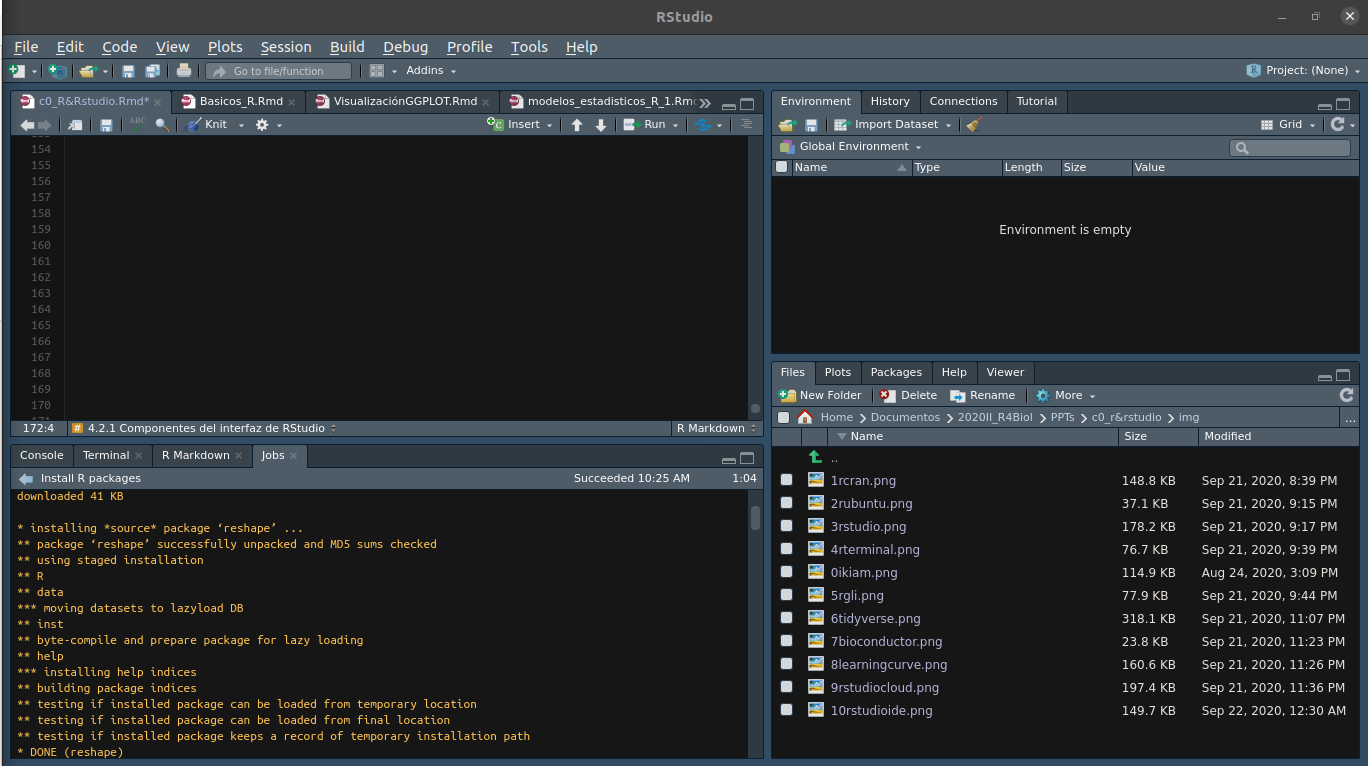


## 4.2 R en RStudio

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:

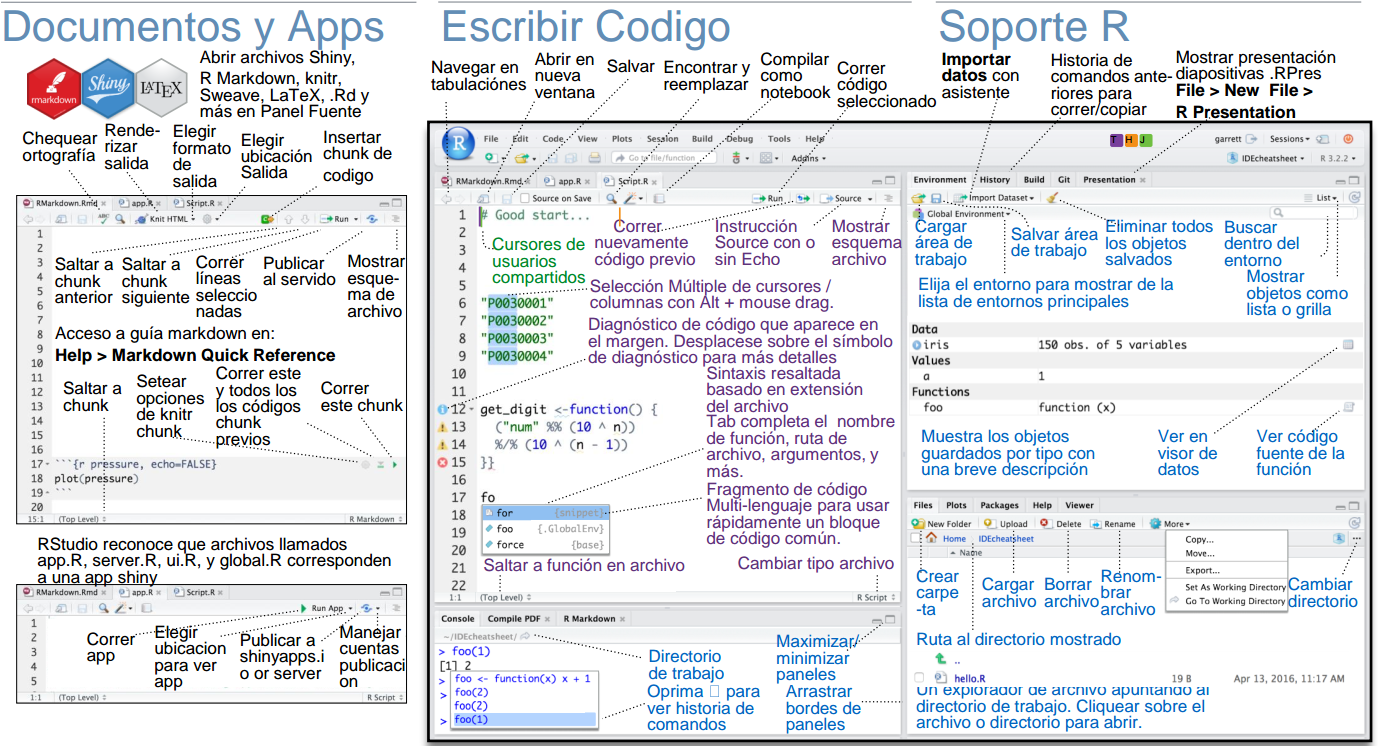


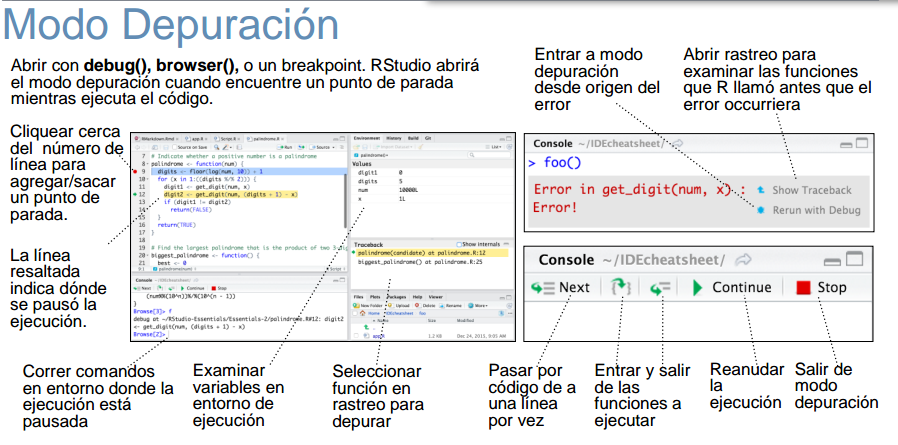
En el interfaz sde RStudio se tiene 4 áreas, además de la barra de opciones en la parte superior.

* **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**;
  1. la pestaña **files** permite ver el historial de archivos trabajados con el programa;
  2. la pestaña **plots** permite visualizar los gráficos que se generen;
  3. la pestaña **packages** permite ver los paquetes descargados y guardados en el disco duro así como gestionar su instalación o actualización;
  4. 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;
  5. 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 señalados a continuación están disponibles en: <https://github.com/rstudio/cheatsheets/blob/master/translations/spanish/rstudio-ide_Spanish_Translation_Monica_Alonso.pdf>





# 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>