

Introducción al diseño y análisis de experimentos
usando R
Agustín Alesso

Índice general

Introducción	5
1. Empezando con R y RStudio	7
1.1. ¿Cómo instalar R y RStudio?	8
1.2. Primera sesión en RStudio	9
2. Introducción a R	15
2.1. ¿Cómo instalar R y RStudio?	16
3. Revisión de estadística básica	23

Introducción

Este pequeño libro tiene como objetivo servir como material de apoyo al curso de Diseños de Experimentos de la Maestría de Cultivos Intensivos de la FCA-UNL. En el se desarrollará

Introducción a R/Rstudio Repaso de estadística básica Principios de diseño Diseños más empleados

Al final del libro se espera que Ud.

- Reconozca la importancia del enfoque estadístico para el diseño de experimentos
- Tenga un sólido entendimiento de los principales métodos utilizados para el análisis de datos provenientes de experimentos diseñados.
- Desarrolle destrezas mínimas para utilizar el paquete estadístico R y el entorno de trabajo RStudio.
- Sepa interpretar los resultados de los análisis estadísticos en contexto de las ciencias biológicas y agrícolas.

Capítulo 1

Empezando con R y RStudio

R es un lenguaje y entorno para el procesamiento, visualización y análisis estadístico de datos. Ha sido creado en 1993 por R. Gentleman y R. Ihaka, ambos científicos del Departamento de Estadística de la Universidad de Auckland (Nueva Zelanda). Actualmente su desarrollo y mantenimiento está a cargo del *R Core Team*. El sitio oficial del proyecto es www.r-project.org

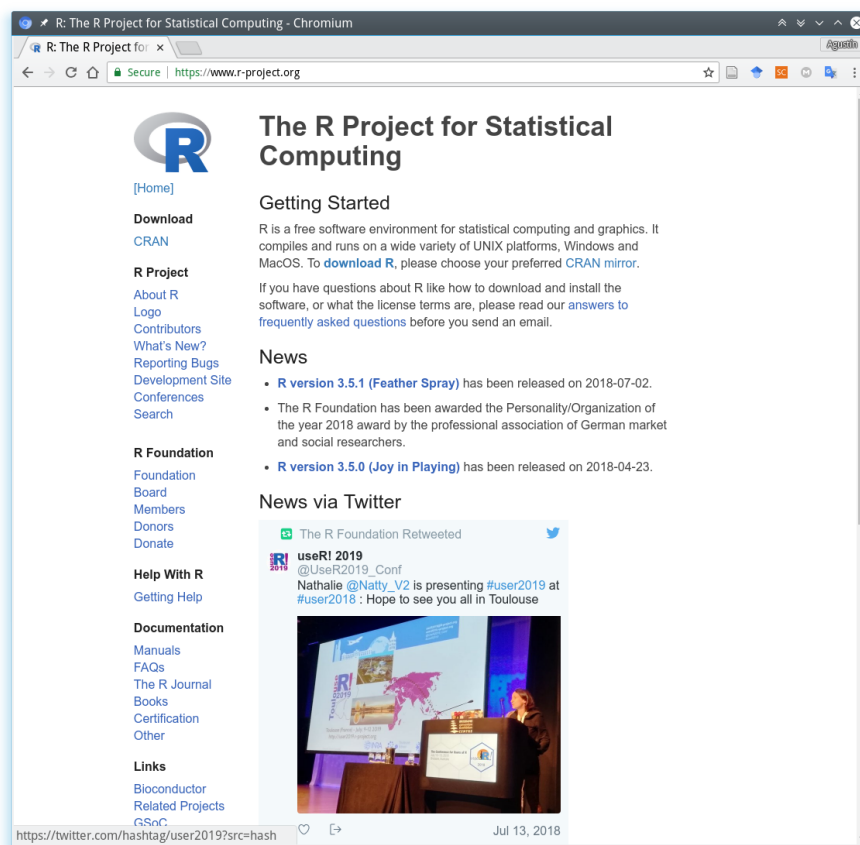
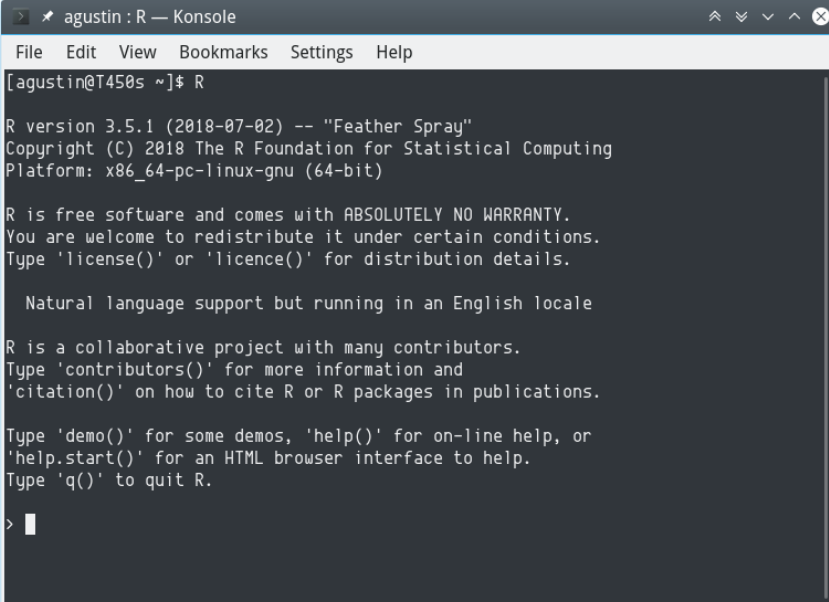


Figura 1.1: Página oficial de R Project

Hoy en día, R es la *lingua franca* del procesamiento y análisis estadístico de datos, tanto en el ámbito académico como comercial dado que es gratuito, multiplataforma, de código abierto (*open source*, liberado con licencia GNU/GPL). Esto lo convierte en un software muy potente y que expresa el estado del arte de los métodos estadísticos ya que la comunidad de usuarios contribuya constantemente con funcionalidades e implementaciones de nuevos métodos y técnicas estadísticas.

Al igual que su antecesor S, la flexibilidad y potencia de R se basa en su interfaz de comandos (CLI, del

inglés *command line interface*) que permite la ejecución de comandos de manera interactiva (en consola) o automática mediante scripts.



```
agustin: R - Konsole
File Edit View Bookmarks Settings Help
[agustin@T450s ~]$ R

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (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.

> 
```

Figura 1.2: Consola o terminal de Windows, Mac OS X y Linux corriendo la última versión estable de R

Existen algunos desarrollos de interfases gráficas (GUIs, del inglés *graphical user interface*), e.g. RCommander, Deducer, que ofrecen la posibilidad de, mediante menues y botones, ejecutar análisis relativamente simples minimizando la necesidad de escribir código.

Por su parte, los entornos de desarrollo integrados (IDE por sus siglas en inglés *integrated development environments*) ofrecen un enfoque intermedio con menús o funciones asistentes que facilitan algunas tareas generales (abrir archivos, carga de datos, exportar gráficos y resultados) pero dejando la codificación del análisis estadístico en manos del usuario mediante la ejecución de scripts. Entre estas alternativas se destaca R Studio (www.rstudio.com) el cual también es de código abierto (licencia GNU/GPL), multiplataforma y ofrece una versión gratuita.

1.1. ¿Cómo instalar R y RStudio?

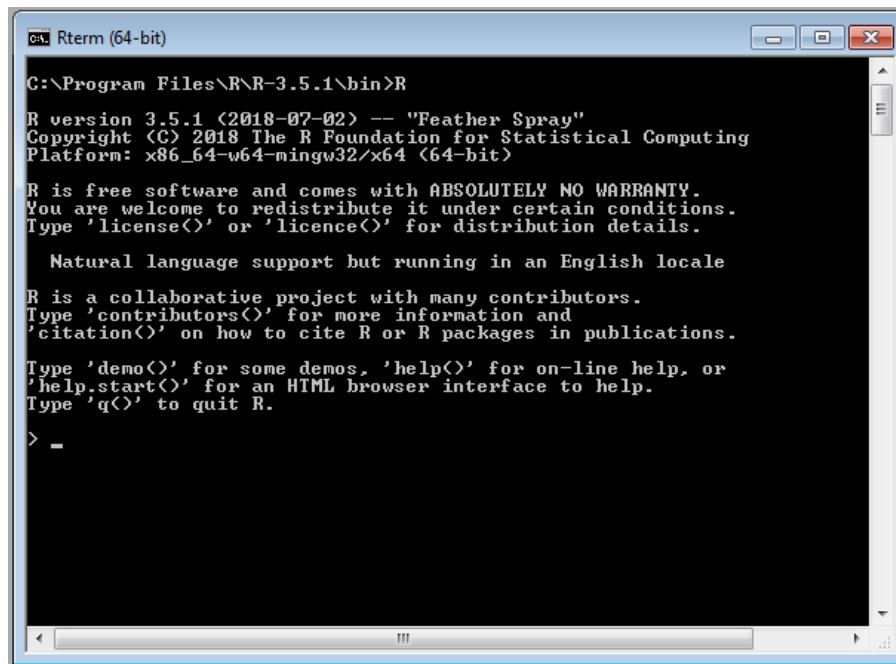
RStudio requiere que el sistema tenga al menos una versión de R instalada. Ambos softwares son multiplataforma y pueden ser ejecutados en sistemas operativos Windows, OS X y Linux. A continuación se describe el procedimiento para instalar R y RStudio bajo Windows.

1.1.1. Instalación de R

- 1) Descargar el archivo instalador correspondiente a la última versión estable de R desde el CRAN¹ (del inglés, *Comprehensive R Archive Network*) visitando el siguiente link ².
- 2) Una vez finalizada la descarga ejecutar el archivo `.exe` y seguir el asistente de instalación con todas las opciones por defecto.

¹CRAN se compone de un conjunto de servidores espejo distribuidos alrededor del mundo que tienen copias de R y sus paquetes. No es necesario escoger el espejo más cercano ya que el espejo nube (<https://cloud.r-project.org>) automáticamente determina de que servidor conviene realizar la descarga.

²Al momento de escribir estas instrucciones la última versión estable de R era la 3.5.1 “Feather Spray”, por lo tanto el link apuntará al archivo `R-3.5.1-win.exe`.



```

C:\Program Files\R\R-3.5.1\bin>R

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 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.

> _

```

Figura 1.3: Consola o terminal de Windows, Mac OS X y Linux corriendo la última versión estable de R

Si la instalación ha sido exitosa el el menú *Inicio > Todos los Programas > R* se encontrarán dos accesos directos R i386 3.5.1 y R x64 3.5.1 los cuales permiten correre la interfase de usuario mínima que viene con la versión de R para Windows.

1.1.2. Instalación de RStudio

- 1) Ir al sitio web de descarga de RStudio: <https://www.rstudio.com/products/rstudio/download/>
- 2) Descargar el archivo de instalación correspondiente a nuestra plataforma o sistema operativo. Por ejemplo: para Windows iniciará la descarga del archivo `RStudio-1.1.453.exe`
- 3) Una vez finalizada la descarga ejecutar el archivo `.exe`³ `RStudio-1.1.453.exe` y seguir el asistente de instalación con todas las opciones por defecto.

Si la instalación ha sido exitosa el el menú *Inicio > Todos los Programas > R Studio* se encontrará el acceso directo a R Studio el cual, mediante el menu contextual (botón derecho del ratón) puede enviarse al Escritorio como acceso directo o bien anclar al menu de Inicio o barra de acceso rápido.

Ahora sí, ya tenemos listo R y RStudio para empezar a trabajar!!

1.2. Primera sesión en RStudio

El entorno de trabajo de RStudio se divide en cuatro paneles: (1) el editor, (2) la consola, (3) entorno, historia de comandos y conexiones y (4) administrador de archivos, gráficos, ayuda y paquetes.

1. **Editor de scripts.** Por defecto este panel no aparece a menos que se cree un nuevo script o se abra uno previamente guardado. Es básicamente un editor de texto plano como el block de notas, aunque tiene algunas funcionalidades importantes:

- Resaltado sintaxis: mediante colores resalta las funciones, variables, comandos o palabras claves del lenguaje **R**

³Al momento de escribir estas instrucciones la última versión estable de R Studio era la 1.1.453 por lo tanto el link apuntará al archivo `RStudio-1.1.453.exe`.

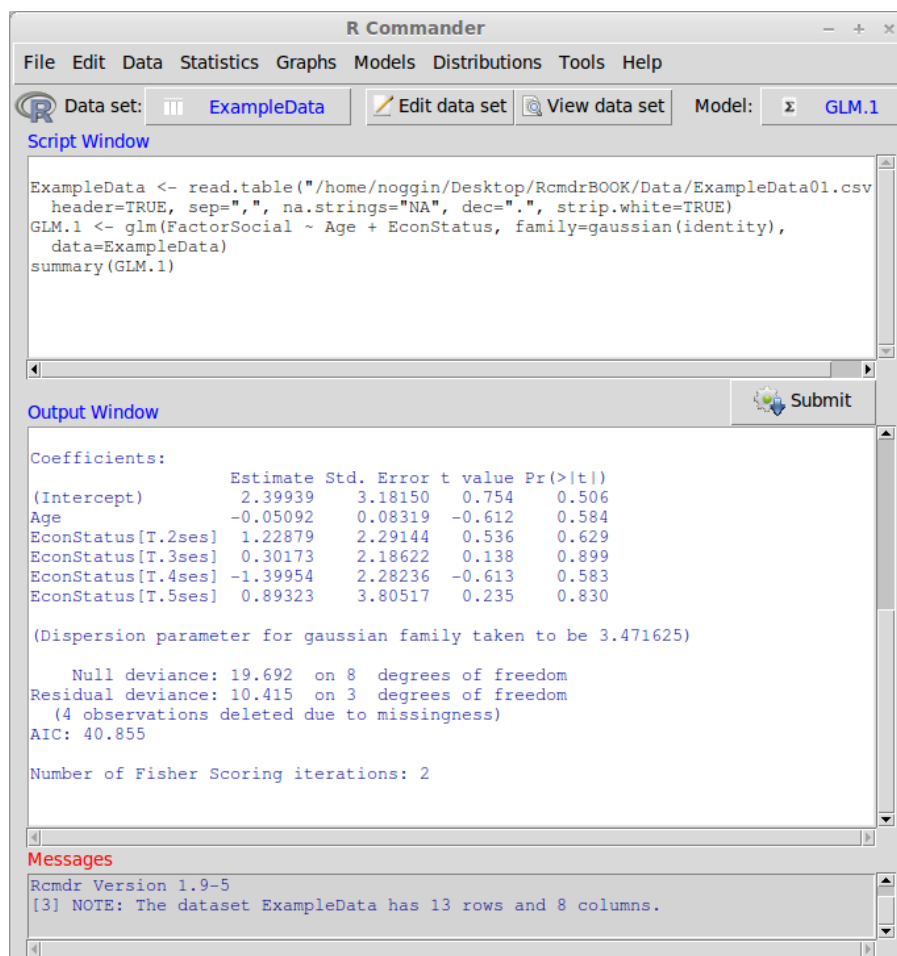


Figura 1.4: Interfase de R Commander

- Sangrado automático: agrega espacios en blanco para mantener la sangría de los bloques de código.
 - Completado automático: muestra sugerencias para completar el comando o argumentos usando la tecla TAB.
2. **Consola.** Es donde reside **R** propiamente dicho. Allí se ejecutan los comandos y se obtienen las salidas de **R**. El símbolo es `>` indica que **R** está disponible para recibir un comando que puede ser tipeado directamente, o bien enviado desde el editor de scripts usando la combinación **CTRL + ENTER** o **CTRL + R**.
 3. **Environmnet/History.** En la primera pestaña se visualizan los objetos (variables, funciones o datos cargados) en el entorno de **R**. En la segunda se puede ver el historial de comandos ingresados o enviados a la consola.
 4. **Files/Plots/Packages/Help/Viewer.** Allí se puede manejar los archivos del directorio de trabajo, visualizar los gráficos generados en **R** con posibilidad de exportarlos en varios formatos, administrar los paquetes o complementos, buscar o explorar el manual de ayuda.

1.2.1. Crear un Proyecto

Antes de comenzar es conveniente crear un proyecto dentro de **RStudio**. Esto permitirá organizar los archivos de datos, las salidas, los scripts, etc., dentro de un directorio de trabajo (*working directory*) y volver a ellos de manera más rápida y eficiente.

1. Ir a **File > New project...** o bien el ícono **New project**.

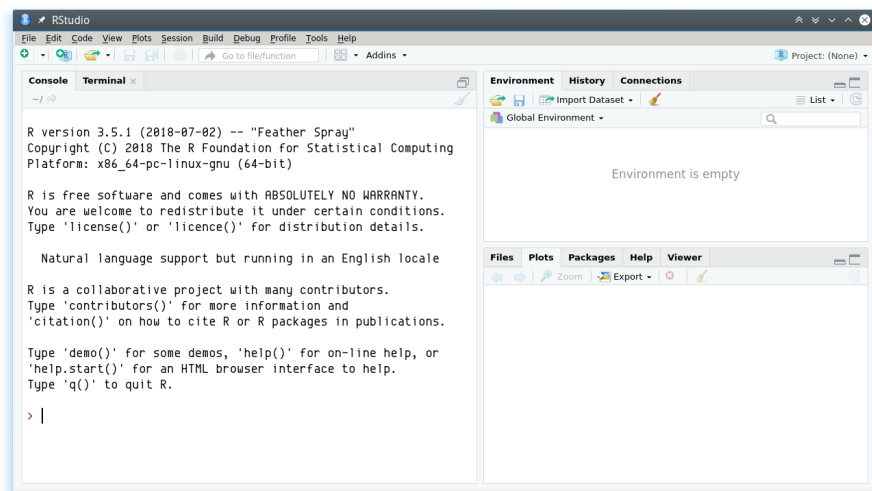


Figura 1.5: Interfase de R Studio

2. Luego seleccionar **New directory** y **Empty project**
3. Una vez en el cuadro de diálogo **Create new project**

En **Directory name** ingresar el nombre del proyecto (e.g. **Diseño2016**) que será a su vez el nombre de la carpeta que **RStudio** va a crear.

Luego en **Create project as a subdirectory of** vamos indicar **donde** queremos que **Rstudio** cree la carpeta.

4. Si todo sale bien, se crea la carpeta con el nombre que indicamos y dentro de ésta un archivo con extensión **.Rproj**

1.2.2. Modo interactivo: la consola

La línea de comando o **consola** es el modo interactivo mediante el cual podemos ejecutar comandos directamente en el intérprete de **R**. El símbolo **>** indica que **R** está disponible esperando una orden. Si la orden no está completa el símbolo se transforma en **+**. Por ejemplo: $2 + 2$

```
2 + 2
```

```
## [1] 4
```

Otro ejemplo: el promedio de los números 1, 3 y 4

```
(1 + 3 + 4) / 3
```

```
## [1] 2.666667
```

El símbolo **#** indica que lo que sigue es un comentario y por lo tanto debe ignorarse

```
# Esto es un comentario
```

1.2.3. Creación de un script

El **Editor de Scripts** (panel 1) es un editor de texto que está conectado con la **consola** y gracias a algunas funcionalidades facilitan la edición de código

Para crear un nuevo script se puede usar uno de los siguientes métodos:

- Ir a al menú **File > New File > R Script**
- Usar el atajo de teclado **CTRL + SHIFT + N**
- Clickear en el primer ícono de la barra de menu

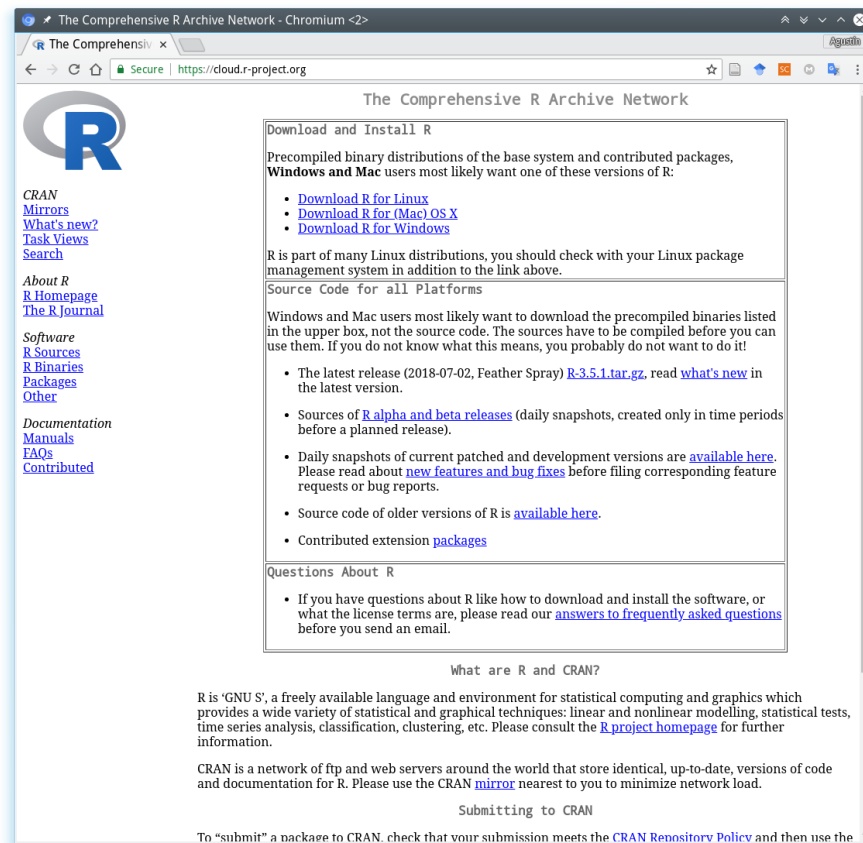
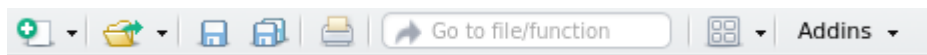


Figura 1.6: Página de descarga de R



Una vez abierto el script en blanco, se pueden empezar a escribir los comandos de **R**, por ejemplo:

```
# Crear un vector con 10 números aleatorios
x <- runif(10, min = 0, max = 10)

# Calcular el promedio de estos números
mean(x)
```

Para ejecutar estos comandos en la consola hay que posicionarse en la línea o seleccionar las líneas que se quieren ejecutar y luego:

- Ir al menú Code > Run Selected Line(s)
- Usar el atajo de teclado CTRL + ENTER o CTRL + R
- Usar el ícono Run de la barra de herramientas de la pestaña del script

Para guardar el script:

- Ir al menú File > Save
- Usar el atajo de teclado CTRL + S
- Usar el ícono con el diskette de la barra de herramientas global o de la pestaña del script activo.

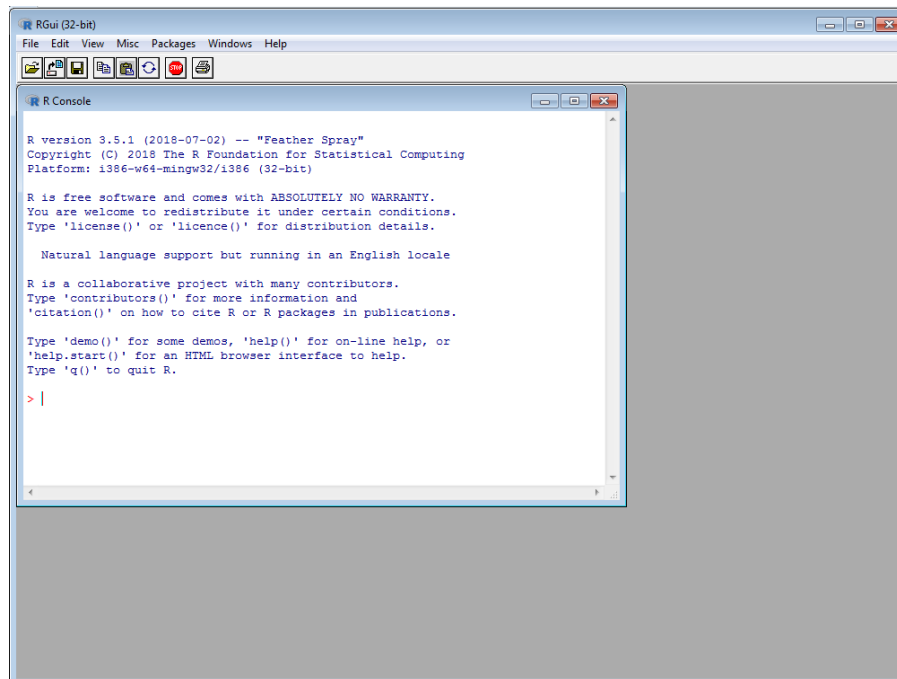


Figura 1.7: R GUI para Windows

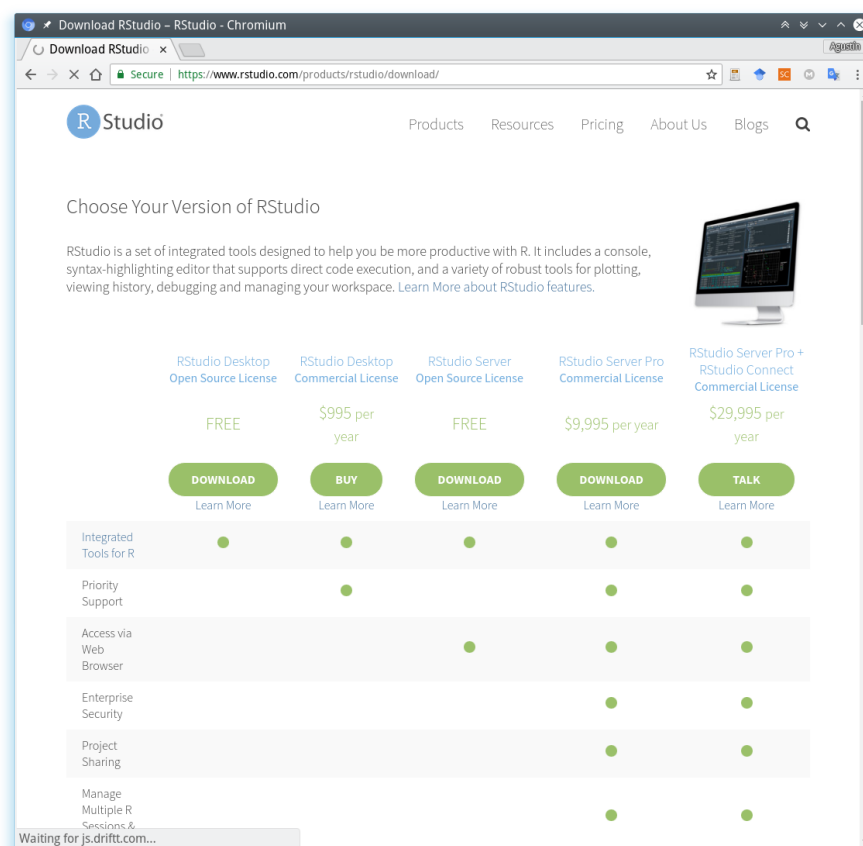


Figura 1.8: Página principal de RStudio

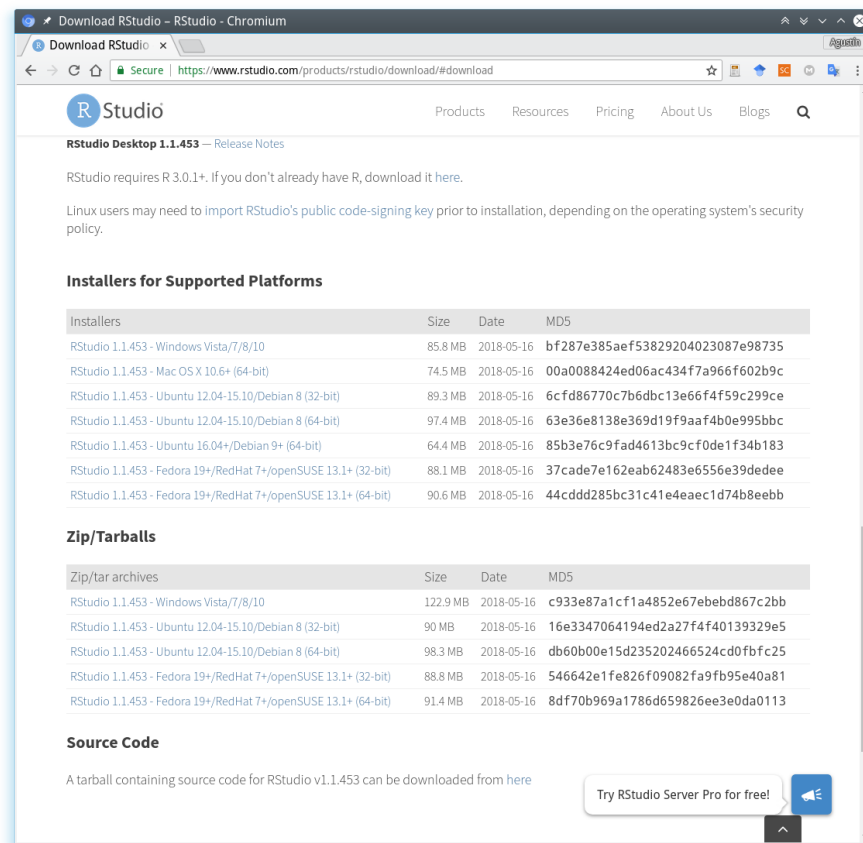


Figura 1.9: Página principal de RStudio

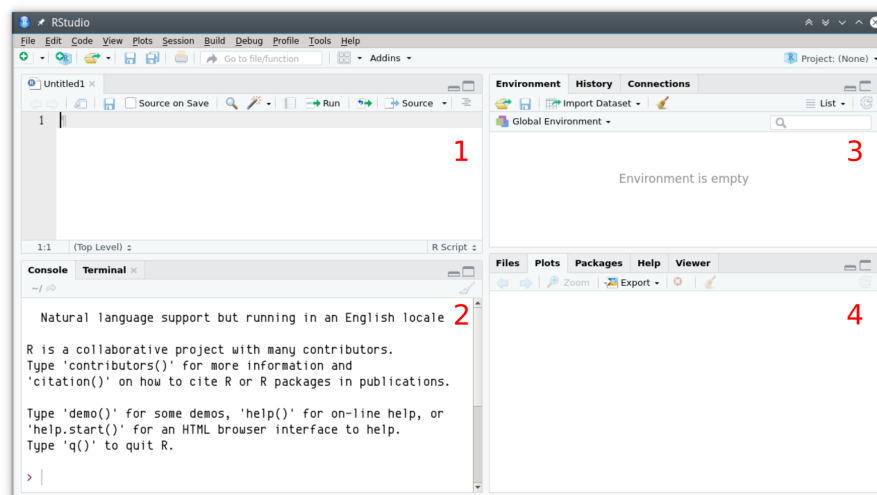


Figura 1.10: Interfase principal de R Studio

Capítulo 2

Introducción a R

R es un lenguaje y entorno para el procesamiento, visualización y análisis estadístico de datos. Ha sido creado en 1993 por R. Gentleman y R. Ihaka, ambos científicos del Departamento de Estadística de la Universidad de Auckland (Nueva Zelanda). Actualmente su desarrollo y mantenimiento está a cargo del *R Core Team*. El sitio oficial del proyecto es www.r-project.org

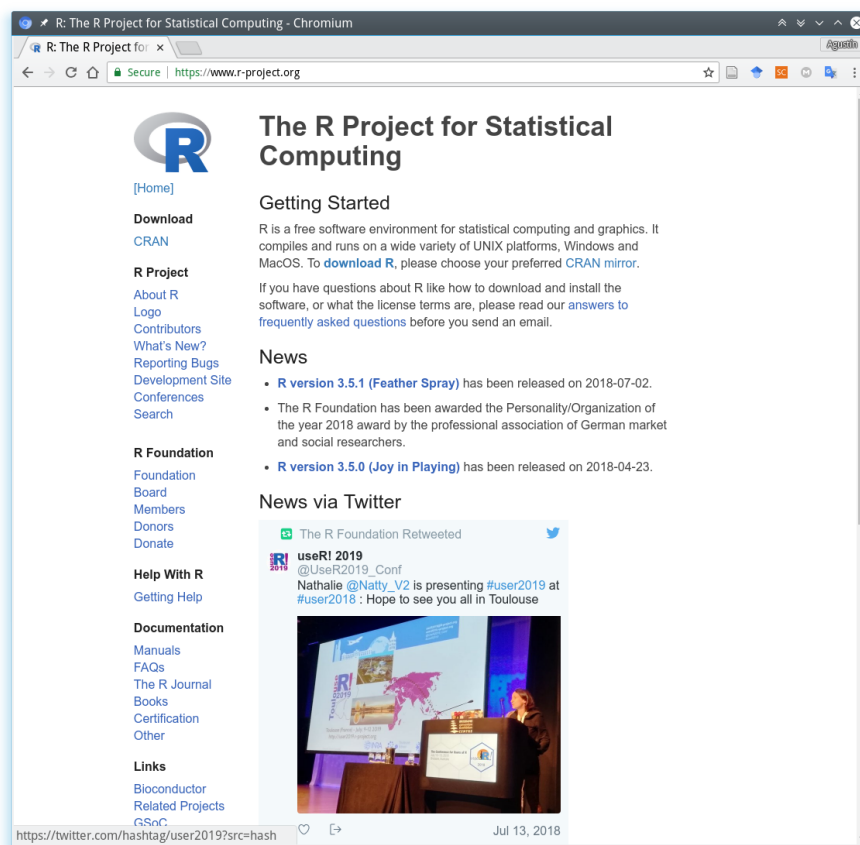


Figura 2.1: Página oficial de R Project

Hoy en día, R es la *lingua franca* del procesamiento y análisis estadístico de datos, tanto en el ámbito académico como comercial dado que es gratuito, multiplataforma, de código abierto (*open source*, liberado con licencia GNU/GPL). Esto lo convierte en un software muy potente y que expresa el estado del arte de los métodos estadísticos ya que la comunidad de usuarios contribuya constantemente con funcionalidades e implementaciones de nuevos métodos y técnicas estadísticas.

Al igual que su antecesor S, la flexibilidad y potencia de R se basa en su interfaz de comandos (CLI, del

inglés *command line interface*) que permite la ejecución de comandos de manera interactiva (en consola) o automática mediante scripts.

```

agustin@T450s ~]$ R

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (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.

>

```

Figura 2.2: Consola o terminal de Windows, Mac OS X y Linux corriendo la última versión estable de R

Existen algunos desarrollos de interfases gráficas (GUIs, del inglés *graphical user interface*), e.g. RCom-mander, Deducer, que ofrecen la posibilidad de, mediante menues y botones, ejecutar análisis relativa-mente simples minimizando la necesidad de escribir código.

Por su parte, los entornos de desarrollo integrados (IDE por sus siglas en inglés *integrated development environments*) ofrecen un enfoque intermedio con menús o funciones asistentes que facilitan algunas ta-reas generales (abrir archivos, carga de datos, exportar gráficos y resultados) pero dejando la codificación del análisis estadístico en manos del usuario mediante la ejecución de scripts. Entre estas alternativas se destaca R Studio (www.rstudio.com) el cual también es de código abierto (licencia GNU/GPL), multiplataforma y ofrece una versión gratuita.

2.1. ¿Cómo instalar R y RStudio?

RStudio requiere que el sistema tenga al menos una versión de R instalada. Ambos softwares son multi-plataforma y pueden ser ejecutados en sistemas operativos Windows, OS X y Linux. A continuación se describe el procedimiento para instalar R y RStudio bajo Windows.

2.1.1. Instalación de R

- 1) Descargar el archivo instalador correspondiente a la última versión estable de R desde el CRAN¹ (del inglés, *Comprehensive R Archive Network*) visitando el siguiente link ².
- 2) Una vez finalizada la descarga ejecutar el archivo `.exe` y seguir el asistente de instalación con todas las opciones por defecto.

¹CRAN se compone de un conjunto de servidores espejo distribuidos alrededor del mundo que tienen copias de R y sus paquetes. No es necesario escojer el espejo más cercano ya que el espejo nube (<https://cloud.r-project.org>) automáticamente determina de que servidor conviene realizar la descarga.

²Al momento de escribir estas instrucciones la última versión estable de R era la 3.5.1 “Feather Spray”, por lo tanto el link apuntará al archivo `R-3.5.1-win.exe`.

Figura 2.3: Consola o terminal de Windows, Mac OS X y Linux corriendo la última versión estable de R

Si la instalación ha sido exitosa el el menú *Inicio > Todos los Programas > R* se encontrarán dos accesos directos R i386 3.5.1 y R x64 3.5.1 los cuales permiten correre la interfase de usuario mínima que viene con la versión de R para Windows.

2.1.2. Instalación de RStudio

- 1) Ir al sitio web de descarga de RStudio: <https://www.rstudio.com/products/rstudio/download/>
- 2) Descargar el archivo de instalación correspondiente a nuestra plataforma o sistema operativo. Por ejemplo: para Windows iniciará la descarga del archivo RStudio-1.1.453.exe
- 3) Una vez finalizada la descarga ejecutar el archivo .exe ³ RStudio-1.1.453.exe y seguir el asistente de instalación con todas las opciones por defecto.

Si la instalación ha sido exitosa el el menú *Inicio > Todos los Programas > R Studio* se encontrará el acceso directo a R Studio el cual, mediante el menu contextual (botón derecho del ratón) puede enviarse al Escritorio como acceso directo o bien anclar al menu de Inicio o barra de acceso rápido.

Ahora sí, ya tenemos listo R y RStudio para empezar a trabajar!!

³Al momento de escribir estas instrucciones la última versión estable de R Studio era la 1.1.453 por lo tanto el link apuntará al archivo RStudio-1.1.453.exe.

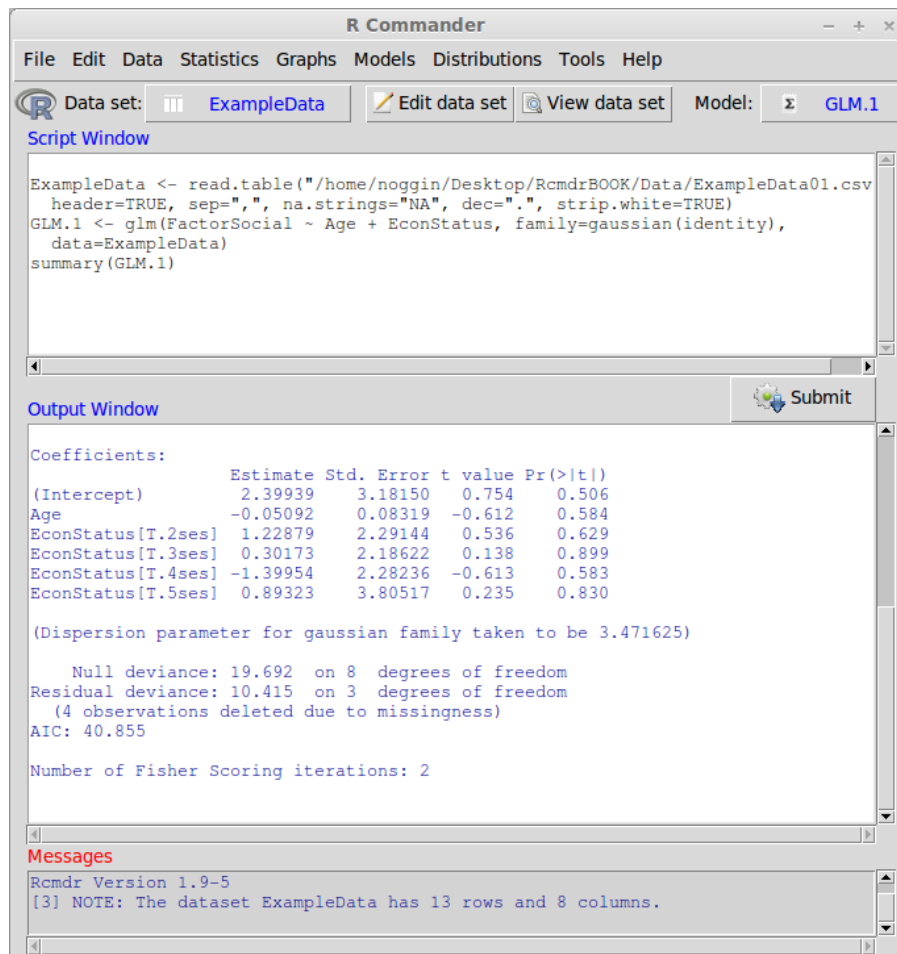


Figura 2.4: Interfase de R Commander

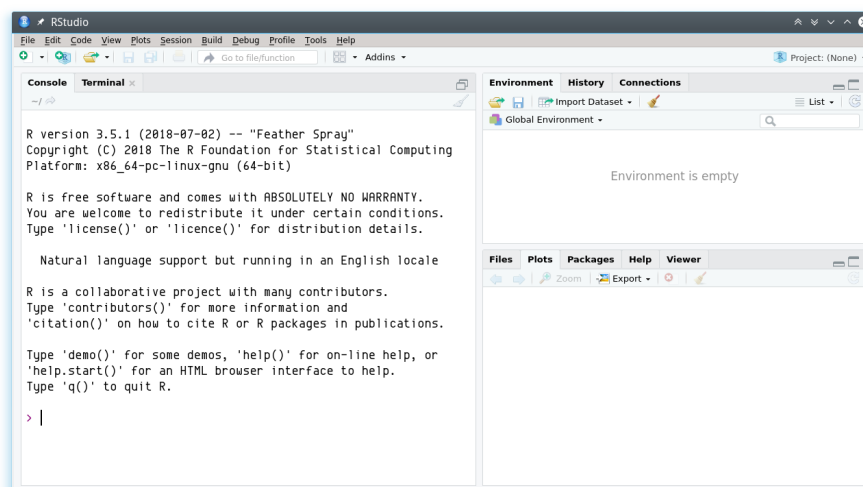


Figura 2.5: Interfase de R Studio

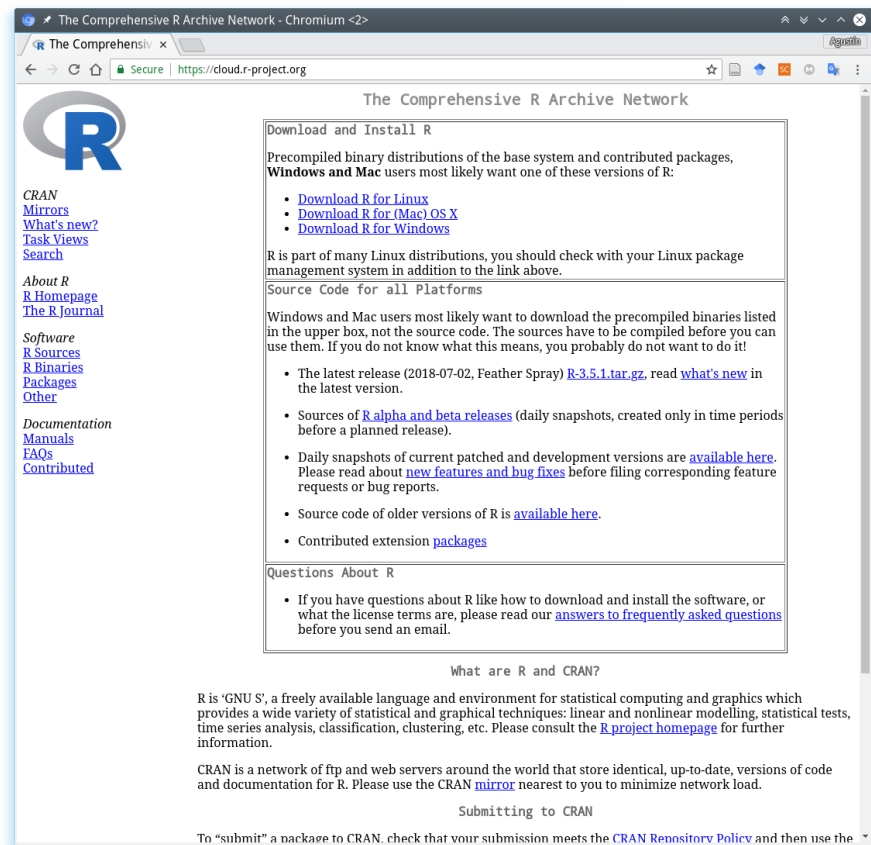


Figura 2.6: Página de descarga de R

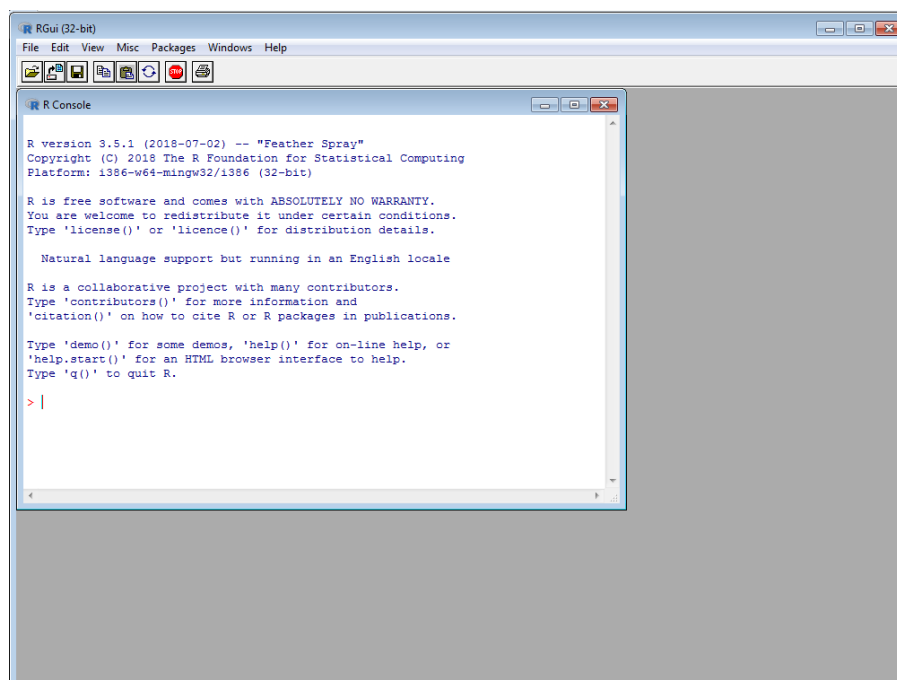


Figura 2.7: R GUI para Windows

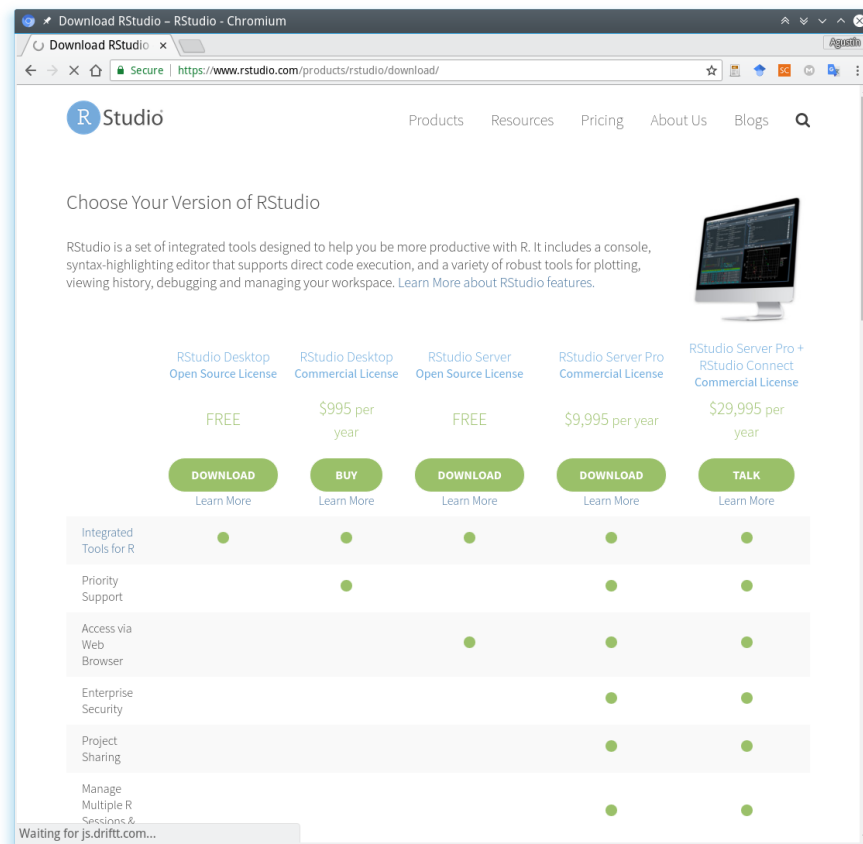


Figura 2.8: Página principal de RStudio

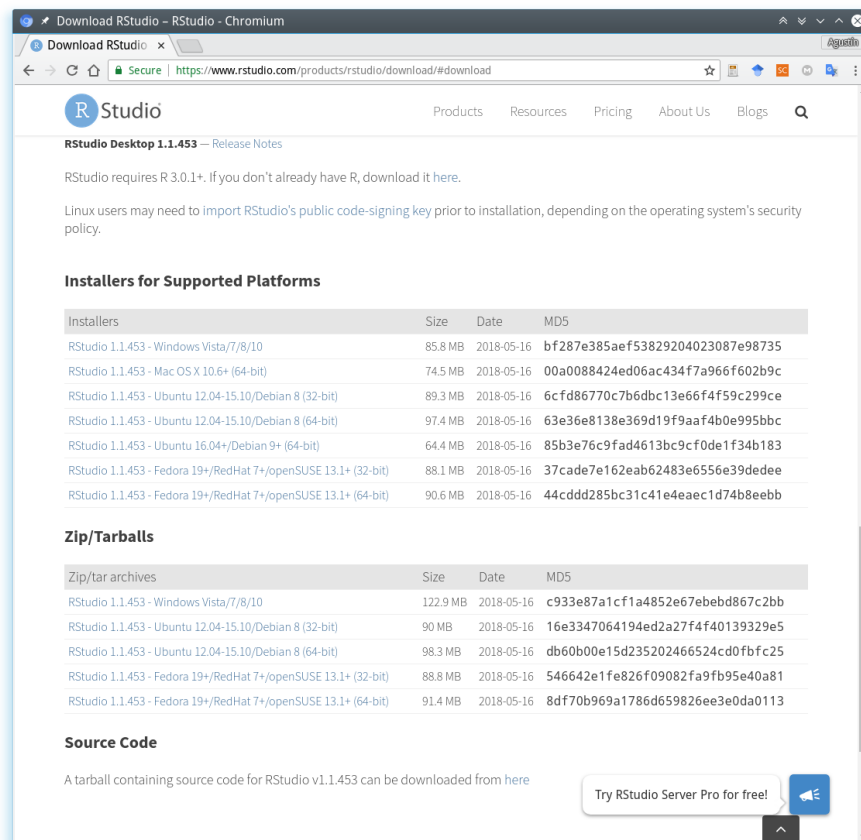


Figura 2.9: Página principal de RStudio

Capítulo 3

Revisión de estadística básica

En este capítulo se revisarán algunos conceptos estadísticos básicos que son fundamentales para el diseño y análisis estadístico de experimentos