



Hablemos de R

By Elsa Toribio

Puntos a Tratar

1. **R.** El lenguaje
 - Qué es R y su Historia
 - Conceptos iniciales
2. **RStudio.** El entorno
 - Dónde descargarlo
 - Características principales
 - La **interfaz** de RStudio
3. **Proyectos de Código abierto de RStudio**
4. **Tipos de datos**
5. **Visualización de Datos**

1

R

El lenguaje



Qué es R!

R es un **lenguaje** de programación, creado por Ross Ihaka y Robert Gentleman, cuya característica principal es que forma un entorno de análisis estadístico para la manipulación de datos, su cálculo y creación de gráficos.





Su Historia



1970

R es la
implementación
moderna de S



1992

En Nueva Zelanda
Ross Ihaka y Robert
Gentleman
Empezaron a desarrollar R



2000

Se lanza la
versión 1.0.0 de
R. Que es una
versión estable



Tras muchos años

...



1993

Primer anuncio
sobre R

*** (S es uno de los
varios lenguajes de
computación
estadística)*



Seguimos hablando de R



Por qué R

Es gratis
Es de software libre
descargable desde
CRAN.



CRAN

Es el “gran almacén”
de R donde están
guardados todos los
paquetes (y sus
versiones)
desarrollados en R con
sus documentaciones.



Entorno

El método más común de
utilizar R es a través
del entorno de
desarrollo RStudio.



Seguimos hablando de R

- La página principal del proyecto “R - project” es

`http : //www.r - project.org,`

en ella podremos conseguir gratuitamente el programa, manuales, paquetes y demás elementos que forman la gran familia que es R.

- R es un proyecto vivo, ya que los usuarios pueden contribuir al proyecto implementando funciones, librerías,... Ningún otro programa estadístico en la actualidad reúne la cantidad de recursos y manejabilidad que posee R.

2

RStudio

El entorno



R vs RStudio dónde descargarlo



R

Es un **lenguaje** de programación

<https://cran.r-project.org/>



R Studio

Entorno de desarrollo integrado
(IDE) para el lenguaje de
programación R

<https://www.rstudio.com/products/rstudio/download/#download>



RStudio Diagrama

Permite

Escribir y editar código



Integra

Sistemas de control de versiones. GIT



Para

Generarlo, verlo y también guardar gráficas



Gestiona

Archivos, objetos y dataframes



La Interfaz de RStudio

Dividida en 4 zonas:

- Consola
- Entorno de variables
- Editor
- Utilidades



RStudio IDE : : GUÍA RÁPIDA



Documentos y apps

RStudio desarrolla herramientas gratuitas y abiertas para R. Su entorno de desarrollo integrado (IDE) facilita el análisis de datos con R.

También ofrece: muchos paquetes R (e.g. tidyverse, sparklyr, ggplot2, dplyr), incluidos **Shiny** (crea aplicaciones web sencillas en R) y **R Markdown** (te permite convertir tus análisis en documentos, informes, presentaciones y paneles de alta calidad; compartirlos y reproducirlos).



Escribe tu código

Abrir en una nueva ventana
Pestaña de navegación
Guardar
Encontrar y reemplazar
Ejecutar el código seleccionado
Ejecutar el código previo
Ejecutar el código completo
Esquema del archivo

Soporte R

Historial de comandos
Importar datos con asistente
Mostrar presentaciones de diapositivas. RPres
Archivo> Nuevo archivo> Presentación R

1. Cursor del usuario
2. Selección: Alt + arrastre del mouse.
3. Resaltado de sintaxis.
4. Autocompletar mediante tabulación: nombres de funciones, rutas de archivos, argumentos y más.
5. Cambia el tipo de archivo
6. Salta a la función en el archivo
7. Directorio de trabajo
8. Presiona ↑ para ver el historial de comandos
9. Arrastra los límites
10. Muestra los objetos guardados por tipo, con una breve descripción
11. Ver en el visor de datos
12. Ver el código fuente de la función
13. Crear una carpeta
14. Subir un archivo
15. Eliminar un archivo
16. Renombrar un archivo
17. Ruta del directorio
18. Buscador de archivos. Haga clic en el nombre del archivo o directorio para abrirlo.

GRÁFICOS

Navegar en los últimos gráficos
Abrir en una ventana
Exportar el gráfico
Eliminar el gráfico
Eliminar todos los gráficos

PAQUETES

Instalar paquetes
Actualizar paquetes
Crear un paquete reproducible desde tu proyecto.
Márcalo para activar el paquete (i.e. library()), o desmarcarlo para desactivarlo (i.e. detach()).
Versión del paquete instalada
Eliminar el paquete.

AYUDA

Página de ayuda con links útiles
Busca dentro de la página de ayuda
Busca una página de ayuda

PANEL DE VISIÓN

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
1	5.1	3.5	1.4	0.2	setosa
2					
3					
4					

1. Filtra las filas por un valor o rango de valores
2. Ordena los valores
3. Busca un valor

1

Abrir en una nueva ventana

Pestaña de navegación

Guardar

Encontrar y reemplazar

Compilar como cuaderno

Ejecutar el código seleccionado

Cursor del usuario

Ejecutar el código previo

Ejecutar el código completo

Esquema del archivo

Selección: Alt + arrastre del mouse.

Código de diagnóstico que aparece en el margen. Pase el mouse sobre los símbolos de diagnóstico para obtener más detalles.

Resaltado de sintaxis.

Autocompletar mediante tabulación: nombres de funciones, rutas de archivos, argumentos y más.

Cambia el tipo de archivo

Saltar a la función en el archivo

Directorio de trabajo

Presiona para ver el historial de comandos

Maximiza/minimiza los paneles

Arrastra los límites

2

1. Editor

Para la ejecución de un script se deberá seleccionar lo que se pretenda ejecutar y hay dos opciones

Run

Cntr + Enter

Panel donde se crean y ejecutan los script

2. Consola

Te permite escribir y ejecutar comandos o instrucciones.

También es el lugar donde obtendrás los outputs de los scripts agregados al editor

3. Entorno de variable

Mostrará aquellas variable las cuales le hayamos asignado un valor y se encuentren cargadas en el entorno.

4. Utilidades

Se muestran las diferentes gráficas que se han cargado en el archivo, además de mostrar nuestro entorno de trabajo.

Historial de comandos

Importar datos con asistente

Mostrar presentaciones de diapositivas .RPres
Archivo> Nuevo archivo> Presentación R

garrett Sessions
IDEcheatsheet R 3.2.2

Environment History Build Git Presentation

Import Dataset

Global Environment

Carga el área de trabajo

Guarda el área de trabajo

Buscador del área de trabajo

Elegir el entorno para mostrar

Mostrar objetos como lista o cuadrícula

3

Data

iris 150 obs. of 5 variables

Values

a 1

Functions

foo function (x)

Muestra los objetos guardados por tipo, con una breve descripción

Ver en el visor de datos

Ver el código fuente de la función

Files Plots Packages Help Viewer

New Folder Upload Delete Rename More

Home IDEcheatsheet

Crear una carpeta

Subir un archivo

Eliminar un archivo

Renombrar un archivo

Cambiar el directorio

Ruta del directorio

hello.R 19 B Apr 13, 2016, 11:17 AM

4

Buscador de archivos. Haga clic en el nombre del archivo o directorio para abrirlo.

3

RStudio

Proyectos de Código Abierto



Proyectos de Código abierto de RStudio

Tidyverse

Tidymodels

Shiny®

R Markdown

Conectividad

Infraestructura R

IDE de RStudio





Proyectos de Código abierto de RStudio

Algunos de los importantes proyectos de código abierto dirigidos o respaldados por RStudio incluyen el siguiente software popular para la ciencia de datos:

- **Tidyverse**

El tidyverse es una colección obstinada de paquetes R diseñados para la ciencia de datos. Todos los paquetes comparten una filosofía de diseño, una gramática y estructuras de datos subyacentes.

El tidyverse consta de 27 paquetes R que incluyen **ggplot2**, dplyr, tidyr y readr.

**** Utilizaremos en este taller en ggplot2, readr ... ****





Proyectos de Código abierto de RStudio

- **Tidymodels**

Tidymodels es una colección cohesiva de paquetes que realizan **tareas** relevantes para el **modelado estadístico** y el **aprendizaje automático**. Los paquetes de Tidymodels comparten una sintaxis para funcionar con los paquetes de Tidyverse.

Actualmente hay 35 paquetes de tidymodels, un aumento de 8 desde 2019. Los más populares: incluyen parsnip, rsample, recipes, tune, y yardstick

- **Shiny®**

Shiny es un paquete R popular y un marco de **aplicación web** que facilita contar historias de datos en aplicaciones web interactivas de apuntar y hacer clic.





Proyectos de Código abierto de RStudio

- **R Markdown**

R Markdown es un formato de **creación de documentos computacionales**, que son informes totalmente reproducibles cuyo análisis se puede volver a ejecutar en nuevos datos con solo hacer clic en un botón. Los documentos de R Markdown se pueden compartir como blocs de notas, presentaciones de diapositivas, páginas web, archivos adjuntos de correo electrónico, documentos impresos y más.

Paquetes populares son `rmarkdown` , `knitr` , `flexdashboard` , `blogdown`, `bookdown`, `destile`, `RTÍCULOS` y `xaringan`.





Proyectos de Código abierto de RStudio

- **Paquetes de conectividad**

RStudio aumenta la eficiencia de los usuarios de R al crear paquetes de R de código abierto que **conectan** a los científicos de datos con hojas de cálculo, bases de datos, marcos de almacenamiento distribuidos para big data, plataformas de aprendizaje automático y los entornos de programación de otros lenguajes, como Python.

Los paquetes de conectividad incluyen:

sparklyr, tensorflow para R,
googlesheets, odbc y reticulate.

- **Herramientas de infraestructura R (r-lib)**

R-lib es una gran colección de paquetes R que facilitan la **creación, búsqueda** y **uso** de herramientas efectivas para el análisis de datos.

Actualmente hay 111 paquetes R-lib. Los más populares son devtools, testthat, roxygen2, pkgdown y usethis.





Proyectos de Código abierto de RStudio

- **Entorno de desarrollo integrado (IDE) de RStudio®**

RStudio es un IDE multilingüe diseñado para Data Science con R y Python. Aumenta la consola de código estándar con un editor que puede

- ❑ mostrar Notebooks
- ❑ iniciar aplicaciones
- ❑ resaltar la sintaxis del código
- ❑ detectar errores de código
- ❑ ejecutar código directamente.

Las herramientas integradas en el IDE también permiten:

- ❑ depurar
- ❑ trazar
- ❑ examinar archivos
- ❑ administrar historiales de proyectos y espacios de trabajo

Juntas, estas herramientas hacen que los científicos y desarrolladores de datos sean mucho más eficientes.



4

RStudio

Tipos de datos



Tipos de datos

Esta sección la
explicaremos con código
en RStudio

1. Introducción de comandos / Tipos de datos
2. Operaciones aritméticas
3. Vectores
4. Matrices
5. DataFrame



Tipos de Datos

Integer	1, 2, 3, 4
Numeric	1, 2.5, 3.6, 4
Character	"El 1º de muchos"
Logical	TRUE/FALSE
Otros tipos	Double (doble precisión) Complex (números complejos) Raw (almacenar bytes)



5

RStudio

Visualización de Datos



Visualización de Datos

1. GRÁFICOS NATIVOS DE R
2. Importar datos de un csv
3. GGPLOT



Funciones Gráficas

`plot()` Función genérica para representar en el plano xy puntos, líneas, etc.

`barplot()` Diagramas de barras.

`piechart()` Diagramas de sectores.

`hist()` Histogramas.

`boxplot()` Diagramas de box-and-whisker.

`stripplot()` Similares a

`boxplot()` con puntos.

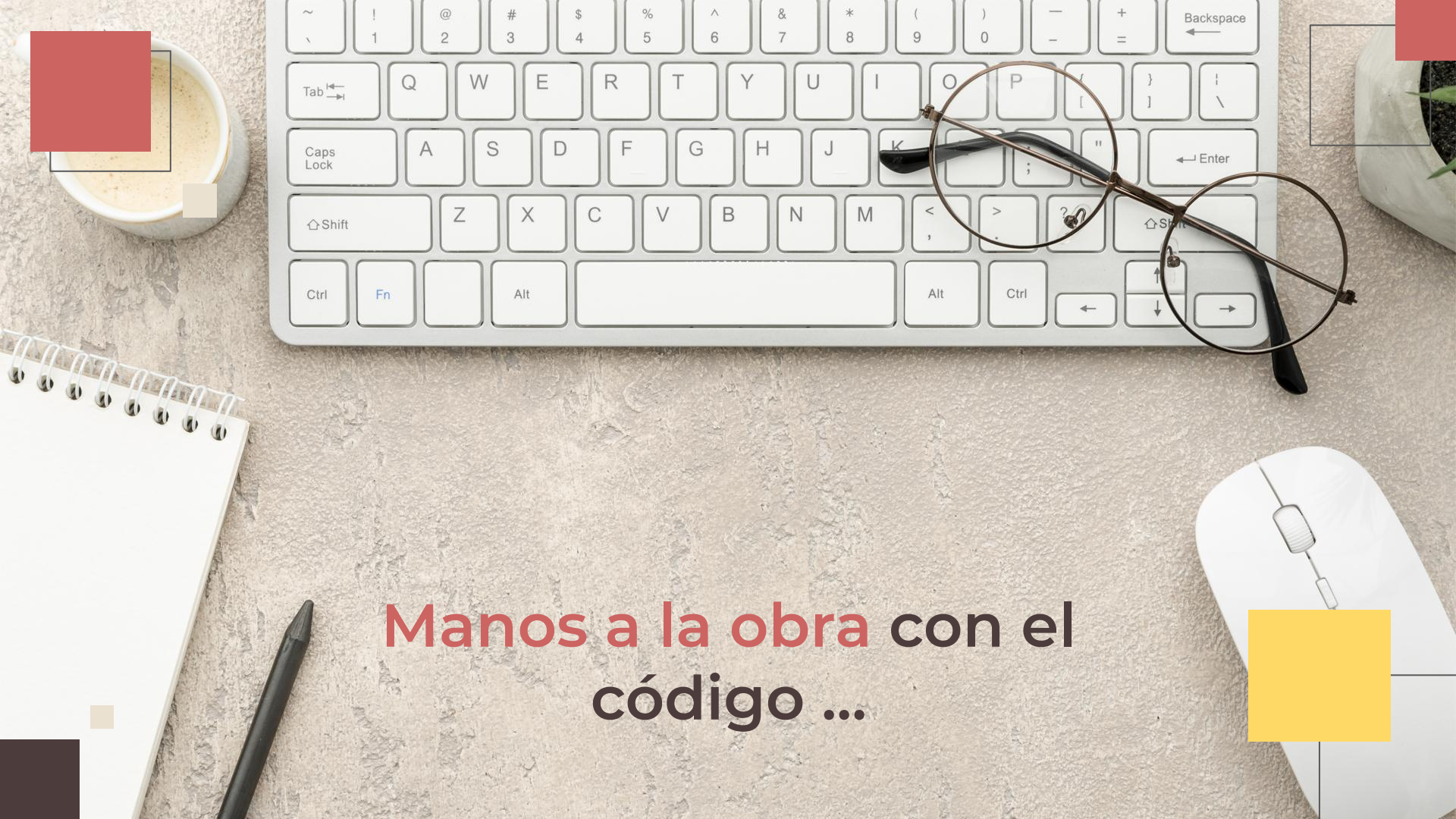
`sunflowerplot()` Representación en el plano xy de diagramas de girasol.

`qqnor()` Diagramas de cuantil a cuantil frente a la distribución normal.

`qqplot()` Diagramas de cuantil a cuantil de dos muestras.

`qqline()` Representa la línea que pasa por el primer y el tercer cuantil.





**Manos a la obra con el
código ...**

Gracias!

Si alguien tiene alguna pregunta?

Mi perfil de linkedin:



<https://www.linkedin.com/in/elsatoribio/>

Os dejo mi correo:

elsatrb@gmail.com

