

Taller de Análisis de datos con R

Estadístico Hugo Andrés Dorado Betancourt
Diciembre 2016

Sesión 1. Introducción a R

Estadístico Hugo Andrés Dorado Betancourt

Contenido

1. Características de R
2. R y R studio.
3. Instalar paquetes y ayudas.
4. Sintaxis
5. Objetos en R

Que es R

- Es un entorno y lenguaje de programación con un enfoque al análisis estadístico.
- Muy popular en el campo de la minería de datos, la investigación biomédica, la bioinformática y las matemáticas financieras.
- Es parte del sistema GNU y se distribuye bajo la licencia GNU GPL. Está disponible para los sistemas operativos Windows, Macintosh, Unix y GNU/Linux.
- Comunidad amplia y creciente de usuarios que aportan y resuelven dudas.

Características de R

- Fue desarrollado inicialmente en el Departamento de Estadística de la Universidad de Auckland (Nueva Zelanda) en 1993.
- Comenzó en los **Bell Laboratories** de AT&T y ahora Alcatel-Lucent en Nueva Jersey con el lenguaje S.
- Se trata de un **lenguaje de programación**, lo que permite que los usuarios lo extiendan definiendo sus **propias funciones** y gran parte de las funciones de R están escritas en el mismo R.
- Para algoritmos computacionalmente exigentes es posible desarrollar bibliotecas en **C, C++ o Fortran** que se cargan dinámicamente.
- Usuarios más avanzados pueden también manipular los objetos de R directamente desde código desarrollado en C

Características de R

- R hereda de S su orientación a objetos.
- Puede integrarse con distintas bases de datos y existen bibliotecas que facilitan su utilización desde lenguajes de programación interpretados como Perl y Python.
- Su capacidad gráfica, permite generar gráficos con alta calidad
- Posee su propio formato para la documentación basado en LaTeX.
- Se ha desarrollado una interfaz, RWeka7 para interactuar con Weka que permite leer y escribir ficheros en el formato arff y enriquecer R con los algoritmos de minería de datos de dicha plataforma.

R y rstudio



Instalación

R

- <https://cran.r-project.org/bin/windows/base/>

Rstudio

- <https://www.rstudio.com/products/rstudio/download2/>

Ayudas en R

- `?rnorm`
- `help.start()`
- `?help.search`
- `help.search("normal")`
- `?apropos`
- `apropos("normal")`
- `?demo`
- `demo(graphics); demo(persp); demo(lm.glm)`

Espacio de trabajo, comentarios

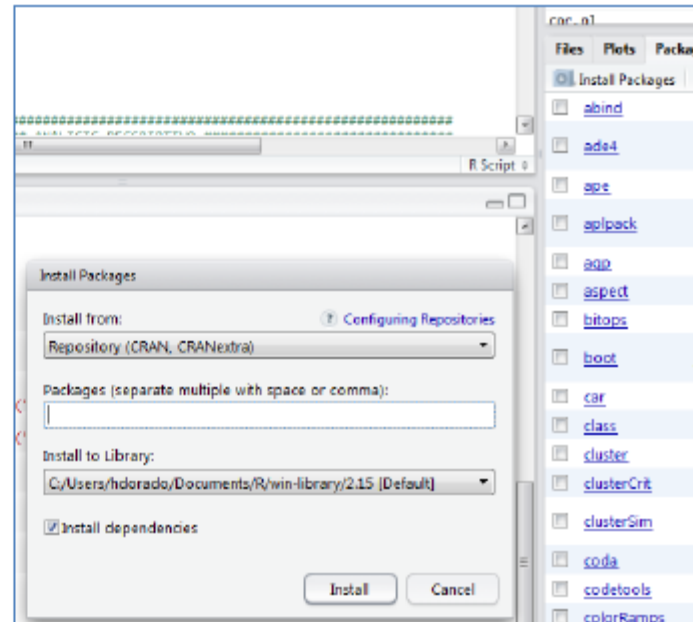
- `Setwd()` # Indicar ruta de trabajo
- `getwd()` # Detectar actual ruta de trabajo
- `Ctrl + Enter`

Instalación de paquetes en R

- `install.packages("")`
- Descargarlos desde CRAN y traerlos en directorios.
- Estar pendiente de las extensiones.

#DESCARGA

```
install.packages("rgdal")  
install.packages("raster")  
install.packages("FactoMineR")  
install.packages("maptools")
```



Instalación de paquetes en R

About R
[What is R?](#)
[Contributors](#)
[Screenshots](#)
[What's new?](#)

Download, Packages
[CRAN](#)

R Project



Colombia

<http://www.lagee.unal.edu.co/CRAN/>

<http://www.icesi.edu.co/CRAN/>



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

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

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

Available Packages

Currently, the CRAN package repository features 38482

[Table of available packages, sorted by date of publication](#)

[Table of available packages, sorted by name](#)

Installation of Packages

Please type `help(*INSTALL*)` or `help("install.packages")` (also contained in the R base source

[CRAN Task Views](#) allow you to browse packages by top available.

Package Check Results

All packages are tested regularly on machines running `Debian`

Objetos

R tiene 5 clases de objetos básicos:

- Carácter
- Numérico
- Entero
- Complejo
- Lógico (True/False)

El objeto base en R es el vector.

Un vector solo contiene objetos de la misma clase.

También existe la posibilidad de crear un objeto lista que puede ser vista como un vector con clases de diferentes.

Se puede crear un objeto vacío con la función `vector()`

Atributos

Los objetos en R tienen atributos

- names, dimnames
- dimensions (matrices, arrays)
- Class
- Length

Los atributos pueden ser accedidos por la función `attributes()`.

Práctica en R

- Entrar datos numéricos y caracteres.
- Evaluación
- Operador “:”
- Mezclar objetos.
- `as.character`, `as.numeric`, `as.character`
- Crear matrices
- Crear matrices a partir de columnas (`cbind`, `rbind`).
- Crear lista.
- Manejo de valores faltantes.
- Crear `data.frame`
- Extraer parte de los objetos.

Resumen

- Aspectos generales de R.
- Manipulación de la interfaz gráfica R studio.
- Instalar paquetes.
- Crear e identificar los diferentes objetos en R.
- Realizar operaciones básicas entre vectores y matrices.