

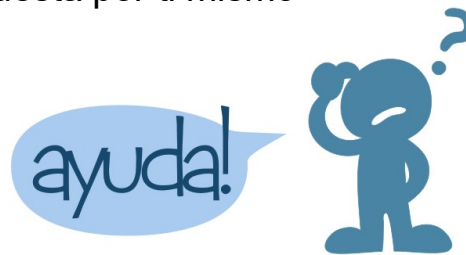
Ciencia de Datos

INTERACCIÓN CON EL MUNDO



Obtener Ayuda

- Preguntar
- Encontrar la respuesta por ti mismo
 - Google
 - Stack Overflow
 - R



Como en VBA también se puede acceder a información del lenguaje y ambiente de programación. (F2)

Algunas funciones importantes de R

Acceder al archivo de ayuda

```
?rnorm
```

Buscar archivos de ayuda

```
help.search("rnorm")
```

Obtener argumentos

```
args("rnorm")
```

<http://cran.r-project.org/doc/contrib/Short-refcard.pdf>

¿Cómo preguntar en Stack OverFlow?

Para preguntar acerca de R

- Que pasos produjeron el problema
- Que esperas obtener
- Que obtienes en vez de eso
- Que versión estás ocupando
- Que sistema operativo usas



¿Cómo preguntar en Stack OverFlow?

Para preguntar acerca de análisis de datos

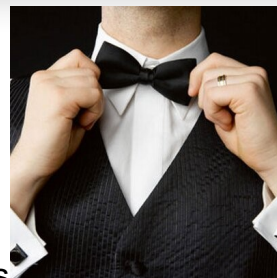
- Que pregunta quieres responder
- Que pasos/herramientas has usado para responderla
- Que esperas obtener
- Que obtienes en vez de eso
- Que alternativas has intentado



Etiqueta de los foros

Lo que se debe hacer:

- Describe tu objetivo
- Se explícito
- Provee la información necesaria
- Se cortés
- Da seguimiento y propón soluciones
- Usa los foros en lugar de un email
- ¡PREGUNTA EN INGLÉS!



Etiqueta de los foros

Lo que se no debe hacer:

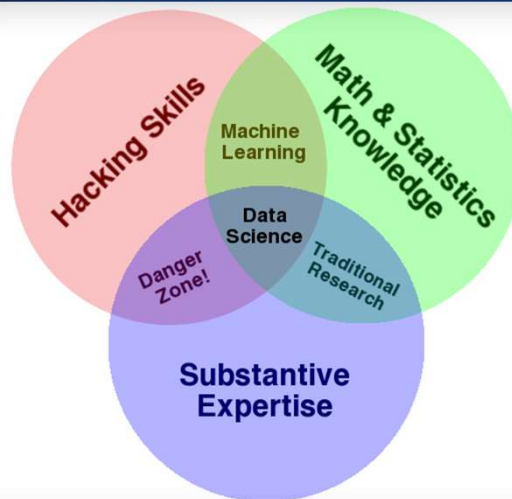
- Asumir automáticamente que hallaste un bug
- Usar los foros como sustituto para hacer el trabajo
- Publicar preguntas de tareas
- Enviar correos en cadenas
- Pedir a otros que arreglen tu código
- Preguntar cuestiones generales de análisis de datos



Se específico en tus preguntas

- Mal
 - HELP! Can't fit linear model!
 - HELP! Don't understand PCA!
- Mejor
 - R 2.15.0 `lm()` function produces seg fault with large data frame, Mac OS X 10.6.3
 - Applied principal component analysis to a matrix - what are U, D, and V^T ?
- Ideal
 - R 2.15.0 `lm()` function on Mac OS X 10.6.3 -- seg fault on large data frame
 - Using principal components to discover common variation in rows of a matrix, should I use U, D or V^T ?

Hacking Skills



- Halla respuestas por ti mismo
- Aprende donde hallar respuestas
- No se intimida por nuevos retos
- No le da miedo aceptar que no sabe algo.
- *Educado pero despiadado.*

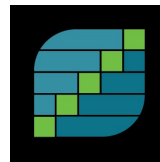
[Google knows it too](#)

Notas para buscar respuestas

- [Stackoverflow](#) (usa la etiqueta "[r]")
- [R mailing list](#) para preguntas de software
- [CrossValidated](#) para preguntas más generales

- Google
 - "[data type] data analysis"
 - "[data type] R package"

 **stackoverflow**



Notas para buscar respuestas

Tipos de datos (data type)

- [Biostatistics](#) médica
- [Data Science](#) web analytics
- [Machine learning](#) computer science/computer vision
- [Natural language processing](#) texts
- [Signal processing](#) electrical signals
- [Business analytics](#) customers
- [Econometrics](#) economic data
- [Statistical process control](#) industrial processes
- etc.

Markdown

- Es un archivo de texto que es formateado en una manera muy específica y que puede ser reconocido por GitHub, R, RStudio.



Markdown básico

```
1 ~~~
2 title: "Sintaxis Basica"
3 author: "Ignacio Trujillo"
4 date: "Sunday, January 03, 2016"
5 output:
6   html_document:
7     theme: cerulean
8 ~~~
9
10 Encabezados
11
12 - # Encabezado Principal
13 - ## Encabezado Secundario
14 - ### Encabezado Terciario
15
16 Listas Desordenadas
17
18 * 1er elemento de la lista
19 * 2do elemento de la lista
20 * 3er elemento de la lista
```

Sintaxis Basica

Ignacio Trujillo

Sunday, January 03, 2016

Encabezados

Encabezado Principal

Encabezado Secundario

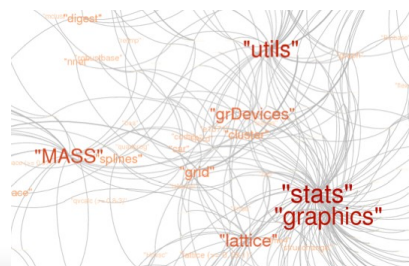
Encabezado Terciario

Listas Desordenadas

- 1er elemento de la lista
- 2do elemento de la lista
- 3er elemento de la lista

Instalación de Paquetes

- Al descargar R de la página de CRAN (*Comprehensive R Archive Network*), se obtiene el sistema base.
- Este sistema tiene las funcionalidades básicas.
- Una razón por la que R es tan útil es porque tiene una gran colección de paquetes que extienden esa funcionalidad.
- Los paquetes se desarrollan y publican por la cada vez mayor comunidad de R.



Obtención de Paquetes

- Para obtener los paquetes la principal fuente será la misma [CRAN](#).
- Se puede obtener información acerca de los paquetes disponibles en CRAN con la función: `available.packages()`.

```
> a<-available.packages()
> head(rownames(a),3)
[1] "A3" "ABCanalysis" "ABCOptim"
> n<-signif(nrow(a),2)
> n
[1] 7500
> |
```

- A inicios del 2016 hay aproximadamente 7500 paquetes en CRAN.
- Continuamente se agregan más y pueden encontrarse ordenados por categoría en la página de [Task Views](#) de CRAN.

Instalación de un paquete

- Se pueden instalar con la función `install.packages()` de R.
- Para instalar un paquete se debe colocar como argumento de esa función.
- Instalemos el paquete *ggplot2*.

```
install.packages("ggplot2")
```

- Para instalar desde otra fuente se tiene que usar `source()`.

```
source("http://bioconductor.org/biocLite.R")
```


Cargar los paquetes

- Se debe tener el paquete instalado.
- El siguiente paso es cargar los paquetes en R, para ello se emplea la función `library()` teniendo como argumento el nombre del paquete.

```
library(ggplot2)
```

- Si un paquete depende de otros para funcionar, los descargará e instalará al momento.

Instalar Rtools

- [Rtools](#) es una colección de herramientas útiles para construir paquetes en R.

Building R for Windows

This document is a collection of resources for building packages for R under Microsoft Windows, or for building R itself (version 1.9.0 or later). The original collection was put together by Prof. Brian Ripley; it is currently being maintained by Duncan Murdoch.

The authoritative source of information for tools to work with the current release of R is the "R Administration and Installation" manual. In particular, please read the "[Windows Toolset](#)" appendix.

Rtools Downloads

Some of the tools are incompatible with obsolete versions of R. We maintain one actively updated version of the tools, and other "frozen" snapshots of them. We recommend that users use the latest release of Rtools with the latest release of R.

The current version of this file is recorded here: [VERSION.txt](#)

Download	R compatibility	Frozen?
Rtools32.exe	R 3.2.x and later	No
Rtools32.exe	R 3.1.x to 3.2.x	Yes
Rtools31.exe	R 3.0.x to 3.1.x	Yes
Rtools30.exe	R 2.15.1 to R 3.0.x	Yes
Rtools212.exe	R 2.14.1 to R 2.15.1	Yes
Rtools214.exe	R 2.13.x or R 2.14.x	Yes
Rtools211.exe	R 2.13.x	Yes
Rtools212.exe	R 2.12.x	Yes
Rtools211.exe	R 2.10.x or R 2.11.x	Yes
Rtools210.exe	R 2.9.x or 2.10.x	Yes
Rtools209.exe	R 2.8.x or R 2.9.x	Yes
Rtools208.exe	R 2.7.x or R 2.8.x	Yes
Rtools207.exe	R 2.6.x or R 2.7.x	Yes
Rtools206.exe	R 2.6.x, R 2.5.x or (untested) earlier	Yes

```
install.packages("devtools")
```