



UNIVERSIDAD  
DE GRANADA



# Sistemas Inteligentes para la Gestión de la Empresa

Sesión 1: Introducción a R

E.T.S. de Ingenierías Informática y de Telecomunicación  
Universidad de Granada

Juan Gómez Romero  
jgomez@decsai.ugr.es

**Departamento de Ciencias de la  
Computación e Inteligencia Artificial**  
<http://decsai.ugr.es>

## **R Project**

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

## **R for Everyone: Advanced Analytics and Graphics** (Jared P. Lander, Addison-Wesley 2013)

<http://proquest.safaribooksonline.com/book/programming/r/9780133257182>

## **R-bloggers**

<https://www.r-bloggers.com/how-to-learn-r-2/>

## **Datacamp Introduction to R** (básico)

[https://www.datacamp.com/courses/free-introduction-to-r/?tap\\_a=5644-dce66f&tap\\_s=10907-287229](https://www.datacamp.com/courses/free-introduction-to-r/?tap_a=5644-dce66f&tap_s=10907-287229)

## **Coursera “Introducción a Data Science: Programación Estadística con R”**

<https://www.coursera.org/learn/intro-data-science-programacion-estadistica-r/home/welcome>

# Herramienta para análisis estadístico, visualización de datos y generación de informes

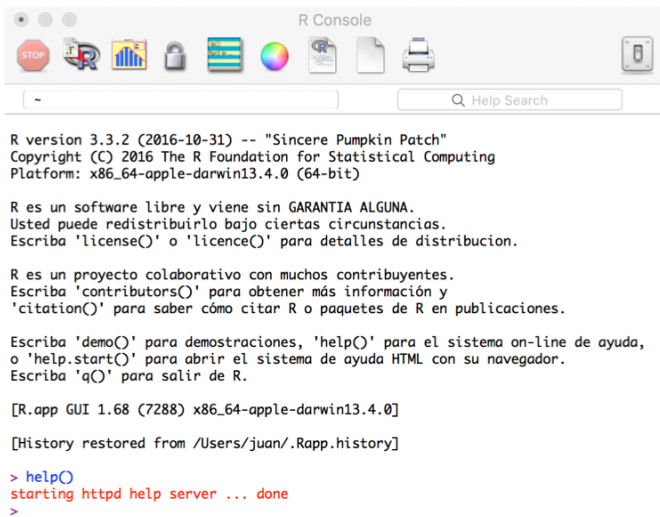
Aplicaciones en Aprendizaje Automático

Componentes:

- "Núcleo": intérprete del lenguaje, paquetes esenciales
- Entorno de programación [RStudio](#): IDE

Extensiones:

- Paquetes adicionales en el repositorio CRAN
- RServer
- Implementaciones alternativas: [Microsoft R](#) (antes Revolution R)
- RStudio notebooks
- Generación de documentos (*knitter*, *swirl*), gráficos (*ggplot2*), interactividad (*shiny*)



```

R Console

R version 3.3.2 (2016-10-31) -- "Sincere Pumpkin Patch"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

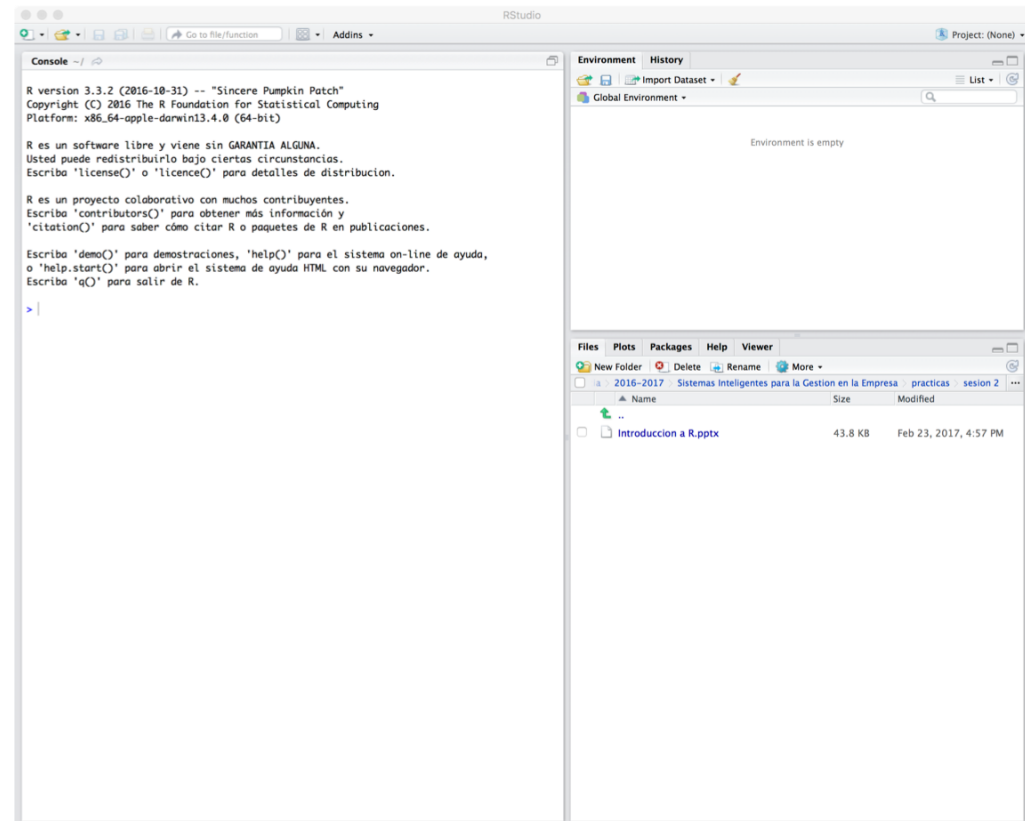
R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.

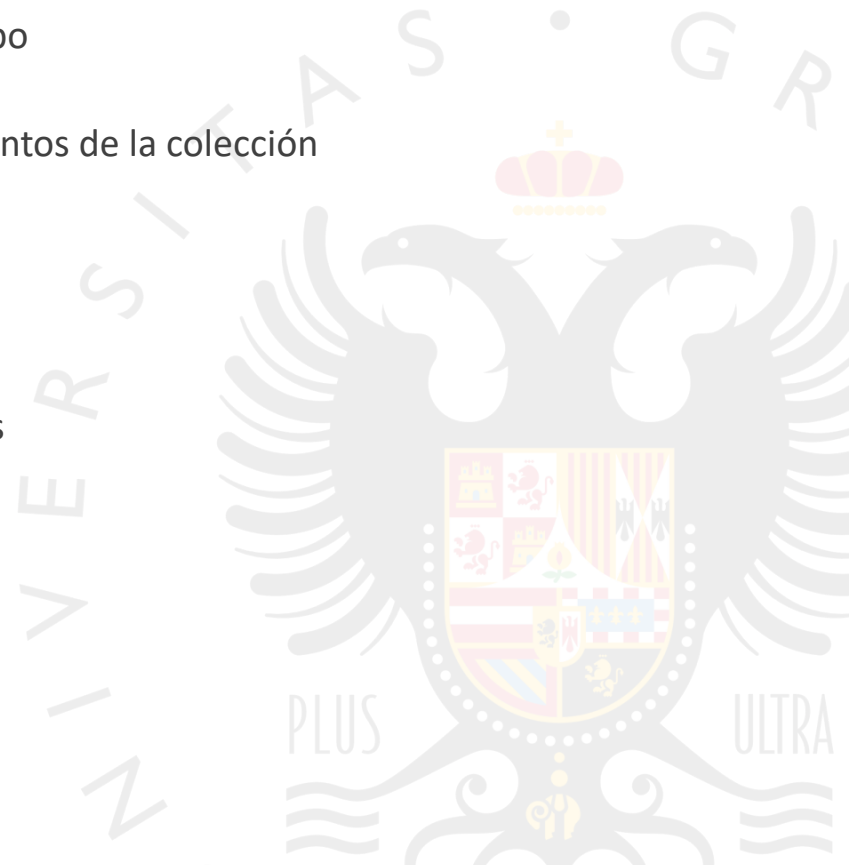
[R.app GUI 1.68 (7288) x86_64-apple-darwin13.4.0]

[History restored from /Users/juan/.Rapp.history]

> help()
starting httpd help server ... done
>
  
```



1. Tipos dinámicos
  - `numeric`, `integer`, `character`, `Date`, `logical`, `factor`
2. Asignación `<-`
3. Colecciones
  - `vector`
    - Colección de elementos, todos del mismo tipo
    - No tienen dimensión
    - Las operaciones se aplican a todos los elementos de la colección
    - Creación: `c()`
    - Rangos de valores: `:`
    - Acceso: `[]` (empezando en 1)
  - `list`
    - Los elementos pueden tener diferentes tipos
    - Creación: `list()`
4. Valores perdidos u omitidos
  - *Not assigned, missing*: `NA`
  - *Nothing*: `NULL`



### 5. Tablas

- `data.frame`
  - Colección de observaciones (filas) de variables (columnas)
  - Cada columna es un vector con un nombre asociado (se puede asignar en la creación del objeto)
  - Las columnas pueden tener tipos diferentes entre sí
  - Creación: `data.frame()`, `read`
  - Funciones: `nrow`, `ncol`, `dim`, `names`, `rownames`
  - Acceso a columnas (pueden utilizarse un vector listando los índices o los nombres que se recuperarán)
    - `[]` (da como resultado un `data.frame`)
    - `$` (da como resultado un vector)
  - Acceso a elementos:
    - `[,]` (da como resultado un vector)
  - Reducción de dimensiones en `data.frame`: parámetro `drop=TRUE` de `[,]`
  - Listado de primeras filas: `head`

### 6. Matrices

- `matrix`
  - Similar a un `data.frame`, todas las columnas tienen el mismo tipo
  - Creación: `matrix()` (especificar valores en primer parámetro, número de filas con `nrow=`; el número de columnas depende de los valores que se asignen)

### .CSV

- **read.table()**, **read.csv()**
  - Da como resultado un `data.frame`

### Otros formatos

- Bases de datos (ODBC)
  - Utilizar paquete RODB: **require(RODB)**
  - Crear conexión: **odbcConnect()**
  - Lanzar consultas: **sqlQuery()** (da como resultado un `data.frame`)
  - Cerrar conexión: **odbcClose()**
  - Sólo una conexión activa
- Otras herramientas
  - SAS, SPSS, Octave
- Tablas HTML
  - Utilizar paquete XML: **require(XML)**
  - Cargar datos: **readHTMLTable()**
- Ficheros binarios
  - Serialización de objetos en R



Soporte nativo para gráficos

- `hist()`, `plot()`, `boxplot()`

Paquete **ggplot2**

- `ggplot(data = <datos>)`
- `+`
- *<especificación de geometría>*
  - `geom_histogram`
  - `geom_density`
  - `geom_point`
- *<especificación de formato, etc.>*
  - `aes()`
  - `labs()`
  - `scale_color_discrete`, `scale_color_continuous`





## Funciones

- **function**
  - No se especifican tipos de los argumentos ni de retorno
  - Pueden definirse argumentos por defecto
  - En la llamada pueden especificarse los argumentos por orden y por asignación

## Control de flujo

- **if() {} else {}**
- **ifelse()**
- **switch()**
  
- **for() {}**
- **while() {}**
  
- **apply()**

Paquete **dplyr** (<http://dplyr.tidyverse.org>)

