



## para Ciencia de Datos: Hoja de Referencia

### R Básico

Aprende R para Ciencia de Datos en [www.datademia.es](http://www.datademia.es)



#### Ayuda

##### Archivos de ayuda

?mean()      Obtén ayuda de una función particular  
help.search('weighted mean')      Busca en los archivos de ayuda una palabra o frase  
help(package = data.table)      Ayuda para un paquete

##### Información sobre un objeto

str(iris)      Obtén un resumen de la estructura de un objeto  
class(iris)      Encuentra la clase a la que pertenece un objeto

#### Librerías

install.packages("data.table")      Descarga e instale un paquete de CRAN  
library(data.table)      Carga el paquete  
lubridate::date      Use una función particular de un paquete  
data(iris)      Carga un conjunto de datos

#### Asignación de variables

> a <- 'apple'      Encuentra el directorio de trabajo actual  
> a      Cambiar el directorio de trabajo actual  
[1] 'apple'

#### Directorio de Trabajo

getwd()      Encuentra el directorio de trabajo actual  
setwd('C://file/path')      Cambiar el directorio de trabajo actual

#### El Entorno

ls()      Muestra todas las variables en el entorno  
rm(x)      Elimina x del entorno  
rm(list = ls())      Eliminar todas las variables del entorno

#### Conversión de tipos

as.logical()	TRUE, FALSE, TRUE	Valores booleanos (VERDADERO o FALSO)
as.numeric()	1, 0, 1	Enteros o decimales
as.character()	'1', '0', '1'	Cadenas de caracteres
as.factor()	'1', '0', '1', levels: '1', '0'	Cadenas de caracteres con niveles. Necesario para algunos modelos estadísticos

#### Comparadores

a == b	Son iguales	a >= b	Mayor que o igual a
a != b	No es igual	a <= b	Menos que o igual a
a > b	Más grande que	is.na(a)	Es NA
a < b	Menos que	is.null(a)	Es nulo

#### Leer y guardar datos

##### Archivo de texto delimitado

df <- read.table('file.txt')  
write.table(df, 'file.txt')

##### Archivo csv (delimitado por comas)

df <- read.csv('file.csv')  
write.csv(df, 'file.csv')

##### Archivo de datos R, un tipo de datos para R

load('file.Rdata')  
save(df, file = 'file.Rdata')

#### Funciones Matemáticas

log(x)	Logaritmo natural	sum(x)	Suma
exp(x)	Exponencial	mean(x)	Media
max(x)	Maximo	median(x)	Mediana
min(x)	Minimo	quantile(x)	Porcentaje de cuantiles
round(x, n)	Redondear	rank(x)	Rango de elementos
signif(x, n)	Redondea a n cifras significativas	var(x)	Varianza
cor(x, y)	Correlación	sd(x)	Desviación estándar

#### Vectores

c(2, 4, 6)	2 4 6	Une elementos en un vector
2:6	2 3 4 5 6	Una secuencia entera
seq(2, 3, by=0.5)	2.0 2.5 3.0	Una secuencia compleja
rep(1:2, times=3)	1 2 1 2 1 2	Repite un vector
rep(1:2, each=3)	1 1 1 2 2 2	Repite elementos de un vector

#### Funciones para Vectores

sort(x)	Devuelve x ordenado.	rev(x)	Devuelve x invertido
table(x)	Ver recuentos de valores	unique(x)	Ver valores únicos

#### Seleccionando Elementos de un vector

##### Por posición

x[4]	El cuarto elemento
x[-4]	Todos menos el cuarto
x[2:4]	Elementos de dos a cuatro
x[-(2:4)]	Todos los elementos excepto dos a cuatro
x[c(1, 5)]	Elementos uno y cinco

##### Por valor

x[x == 10]	Elementos que son iguales a 10
x[x < 0]	Todos los elementos menores que cero
x[x %in% c(1, 2, 5)]	Elementos en el conjunto 1, 2, 5

##### Por nombre

x['apple']	Elemento con el nombre "manzana"
------------	----------------------------------

#### Matrices

m <- matrix(x, nrow = 3, ncol = 3)      Crea una matriz de x

	m[2, ]	Selecciona una fila
	m[, 1]	Selecciona una columna
	m[2, 3]	Selecciona un elemento

#### Listas

l <- list(x = 1:5, y = c('a', 'b'))      Una lista es una colección de elementos que pueden ser de diferentes tipos

l[[2]]	Segundo elemento de l
l[1]	Nueva lista con solo el primer elemento
l\$x	Elemento llamado x
l['y']	Nueva lista con solo un elemento llamado y

#### Data Frames

df <- data.frame(x = 1:3, y = c('a', 'b', 'c'))

x	y
1	a
2	b
3	c

View(df) - Ver el data frame entero  
head(df) - Ver las primeras 6 filas

##### Seleccionando como lista

df\$x     

df[[2]]     

nrow(df) - Número de filas  
ncol(df) - Número de columnas  
dim(df) - Número de filas y columnas

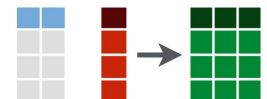
##### Seleccionando como matriz

df[2, ]     

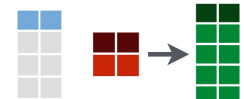
df[, 2]     

df[2, 2]     

##### Cbind - juntar columnas



##### Rbind - juntar filas



Aprende R para Ciencia de Datos en [www.datademia.es](http://www.datademia.es)