



R

Lenguaje de programación creado por Ross Ihaka y Robert Gentleman, pensado para la manipulación, análisis estadístico y visualización de datos.

Archivos con extensión .R

Con él podemos editar en el Editor de Código de RStudio o ejecutar desde consola.

Archivos con extensión .RData

Almacena todos los objetos que están cargadas en memoria.

Archivos con extensión .Rmd

Almacena un proyecto de R Markdown.

Archivos con extensión .Rproj

Proyectos de RStudio que engloban un directorio de trabajo con su espacio de trabajo, historial y documentos de origen.

Array

Vector multidimensional o una matriz multidimensionales donde todos los elementos son del mismo tipo.

Consola

Muestra los resultados de todas las operaciones ejecutadas.

Dataframe

Similar a una matriz, puede contener datos elementos de distinto tipo.

data.frame ()

Función para construir un dataframe.

cbind()

Función para combinar dataframe a través de columnas.

dim()

Devuelve la dimensión del dataframe (filas por columnas).

head()

Función que muestra la cabecera (primeros 6 registros) del dataframe.

ncol()

Devuelve el número de columnas (variables) del dataframe.

nrow()

Devuelve el número de filas (observaciones) del dataframe.

rbind()

Función para combinar dataframe a través de filas.

str()

Devuelve la estructura interna de cualquier objeto en R.

str()

Devuelve la estructura del dataframe.

tail()

Función que muestra los últimos 6 registros del dataframe.

Date

Clase de objeto para trabajar con fechas.

as.character()

Función para convertir variables de clase Date en variables de clase character.



<code>format()</code>	Función para convertir variables de clase Date en variables de clase character.
<code>diff()</code>	Calcula la diferencia en días, entre los términos sucesivos de un vector de fechas declaradas con la clase Date.
<code>unclass()</code>	Muestra la representación interna de un objeto en R.
<code>Sys.Date()</code>	Lee la fecha actual del reloj de nuestro ordenador y la codifica como un objeto de la clase Date.
<code>as.Date()</code>	Convierte un dato de fecha en modo carácter y a la clase Date. Con el argumento <i>format</i> podemos especificar de qué forma ha sido introducida la fecha.
<code>difftime()</code>	Calcula diferencia de fechas.
Directorio de trabajo	Lugar de donde se leen y guardan los archivos que utilicemos en la sesión.
<code>getwd()</code> get working directory	Comando para obtener cuál es el directorio de trabajo actual
<code>setwd()</code>	Función utilizada para cambiar el directorio de trabajo.
Editor de código	Aquí se escribe, se edita o pega el script para luego ser ejecutado.
Environment (Entorno)	Muestra todas las variables que tenemos cargadas en memoria.
Estructuras de control	Permiten controlar el flujo de ejecución de nuestra secuencia de comandos.
<code>break()</code>	Produce la salida inmediata del bucle.
Bucle	Composición compuesta que se repite n veces hasta alcanzar un límite definido.
<code>for</code>	Comando para implementar bucles. <code>for(i in secuencia){ expresión }</code>
<code>if</code>	Para condicionar la ejecución u omisión de una instrucción. La condición debe ser una expresión del tipo booleana y siempre debe ir dentro de un paréntesis. A continuación, debemos escribir el proceso (instrucción compuesta) que queremos que ocurra en caso de que la sentencia utilizada como condición sea verdadera. Agregamos un else, y finalmente el proceso a utilizarse cuando la sentencia es falsa. <code>if(condición){ proceso 1 } else { proceso 2 }</code> Si la condición es verdadera : <code>if(condición){ proceso }</code>
<code>next()</code>	Produce un salto inmediato a la siguiente iteración.



repart()	Comando para realizar bucles que repite un conjunto de instrucciones hasta que se satisfaga un criterio de parada. repeat { expresión if (condición) break }
while	Comando que realiza un ciclo pero no conocemos el número de ciclos que se precisan. while(condición){ expresión }
Función	Toma datos de entrada y los usa para computar un resultado/producto. Tienen un nombre que utilizaremos para llamarla y entre paréntesis escribiremos los argumentos o parámetros que la función necesita separados por comas.
arrange()	Reordena filas de un conjunto de datos.
mutate()	Añade nuevas variables/columnas o transforma variables existentes.
read.table(), read.csv() y read.csv2()	Permite importar y leer archivos.csv
aggregate()	Divide los datos en subconjuntos, calcula estadísticas de resumen para cada subconjunto y devuelve el resultado en un grupo por formulario.
as.character()	Convierte a tipo carácter.
as.factor()	Convierte a tipo factor.
as.integer()	Convierte a tipo entero.
as.logical()	Convierte a tipo lógico o booleano.
as.numeric()	Convierte a tipo numérico.
as.ordered()	Convierte a tipo factor ordenado.
class()	Función para conocer la clase de objeto.
addmargins()	Agrega las sumatorias marginales, de fila y columna a una tabla de contingencia.
cor()	Devuelve el coeficiente de correlación.
cor.test()	Pruebe la asociación entre muestras pareadas, utilizando uno de los coeficientes de correlación.
prop.table()	Tabla de contingencia con valores porcentuales.
round()	Función para redondear números.



shapiro.test()	Función para calcular la prueba de normalidad.
t.test()	Función para prueba de la media de una población normal.
table()	Función para crear tabla de contingencia.
var.test()	Prueba F para comparar dos varianzas.
filter()	Devuelve un conjunto de filas según una o varias condiciones lógicas.
group_by()	Agrupa un conjunto de filas seleccionado, en un conjunto de filas de resumen de acuerdo con los valores de una o más columnas o expresiones.
ordered()	Ordena una variable categórica.
print()	Imprimimos el resultado en la consola.
read.table()	Permite leer tablas en general.
read.table(x,dec...)	El character que se utilizó como decimal, suele ser el punto o la coma.
read.table(x, fill...)	TRUE/FALSE. Para indicar si se deben incorporar las filas incompletas. Por defecto es TRUE.
read.table(x, header...)	TRUE/FALSE Sirve para indicar si la tabla comienza con los nombres de las columnas.
read.table(x, na.strings)	Cómo se indican los datos faltantes en la tabla, por default es "NA"
read.table(x, nrow...)	Por defecto es -1, porque asume que la cabecera ocupa solo una fila.
row.names()	Puede ser un vector que dé los nombres de las filas, o un número que dé la columna de la tabla que contiene los nombres de las filas, o un string de caracteres que dé el nombre de la columna de la tabla que contiene los nombres de las filas.
read_excel(), read_xls() y read_xlsx()	Permite importar y leer archivos .xls, .xlsx
col_names	TRUE para usar la primera fila como nombre de las columnas.
read_excel(), read_xls() y read_xlsx() na	Cómo se indican los datos faltantes en la tabla, por default es "" (vacío).



<code>read_excel()</code> , <code>read_xls()</code> y <code>read_xlsx()</code> sheet	En caso de querer leer otras pestañas que no sea la primera, debemos especificarlo.
<code>read_excel()</code> , <code>read_xls()</code> y <code>read_xlsx()</code> skip	En caso que tengamos una tabla cuya cabecera sea mayor a 1, debemos corregir este parámetro con números positivos.
<code>rename()</code>	Renombra variables en un conjunto de datos.
<code>rep()</code>	Función que permite crear un vector del largo que queramos repitiendo el número o carácter que le indiquemos.
<code>select()</code>	Devuelve un conjunto de columnas (variables).
<code>seq()</code>	Función para escribir una secuencia de números.
<code>set.seed()</code>	Función para generar datos aleatorios.
<code>subset()</code>	Genera otro dataframe seleccionando un subconjunto de datos a partir de uno original.
<code>summarise()</code>	Funciona de forma análoga a la función <code>mutate</code> , excepto que en lugar de añadir nuevas columnas crea un nuevo data frame.
<code>summarise() / summarize()</code>	Genera resúmenes estadísticos de diferentes variables en el conjunto de datos.
<code>tapply()</code>	Función que hace los cálculos por columna del dataframe y por las categorías que defina el usuario, dentro del grupo de observaciones.
<code>as_tibble()</code>	Función que sirve para forzar los objetos <code>data.frame</code> a <code>tbl_df</code> .
<code>tibbles (tbl)</code>	Similares a <code>data.frames</code> pero presentan ciertas mejoras que facilitan su manejo.
<code>write.table()</code>	Genera un archivo con los datos almacenados en un objeto, típicamente un <code>dataframe</code> , aunque puede ser de cualquier otro (vector, matriz, etc)
History (Historial)	Vemos el historial de las órdenes ejecutadas en consola.
<code>list()</code>	Función para crear una lista.
<code>Lista</code>	Colección ordenada de otros objetos, llamados componentes.
<code>length()</code>	Devuelve el número de elementos de un vector o una lista.
<code>names()</code>	Función que devuelve un vector con los nombres de los componentes.



Matriz Vector bidimensional donde todos los elementos son del mismo tipo.

`cbind()` Añade una columna a la matriz en el orden que escribamos la función.

`colnames()` Devuelve el nombre de las columnas de una matriz.

`dim()` Permite crear una matriz a partir de un vector estableciendo su dimensión.

`dim()` Devuelve las dimensiones de una matriz.

`dimnames()` Devuelve el nombre de las dimensiones de una matriz.

`length()` Devuelve el número total de elementos de una matriz.

`matrix()` Función para crear una matriz.

`mode()` Devuelve el tipo de datos de los elementos de una matriz.

`ncol` Argumento que define el número de columnas de una matriz.

`nrow` Argumento que define el número de filas de una matriz.

`rbind()` Añade una fila a la matriz.

`rownames()` Devuelve el nombre de las filas de una matriz.

Paquete Colección de funciones R, datos y código compilado.

Permite instalar un paquete(el nombre entre comillas).

`install.packages("package_name")`

Permite ver qué paquetes están instalados.

`library()` Función que activa paquetes.

`remove.packages()` Permite eliminar paquetes.

`update.packages()` Permite actualizar paquetes.

Paquete dplyr

Contiene una colección de funciones para realizar operaciones comunes de manipulación de datos.

Paqueteforcats

Es parte del ecosistema Tidyverse pensado para trabajar con factores.

`fct_c()` Concatena factores combinando niveles.



fct_collapse()	Función que agrupa los niveles a grupos que serían las nuevas etiquetas del nivel.
fct_count()	Muestra listado de frecuencia de los niveles de la variable.
fct_drop()	Elimina los niveles que no se utilizan.
fct_expand()	Incorpora niveles a la lista de niveles de un factor.
fct_explicit_na()	Función que etiqueta valores NA para visualizarlos.
fct_other()	Agrupa en “otras/os”, varias categorías de una variable con poca frecuencia.
fct_recode()	Función que permite recodificar los niveles de la variable.
fct_relevel()	Función que ordena variable categórica.
Paquete lubridate	Dispone de diversas funciones que facilitan la extracción de componentes de un objeto fecha y hora de clase POSIXct.
ceiling_date()	Redondea la fecha por exceso en la unidad especificada .
dmy_hms()	Incorpora la fecha en el orden día-mes-año-horas-minutos-segundos sin importar los separadores a usar.
floor_date()	Redondea la fecha por defecto en la unidad especificada.
now()	Función que indica hora actual.
round_date()	Redondea la fecha en la unidad que especifiquemos.
today()	Función que indica día actual.
ymd_h()	Fecha en el orden año-mes-día-hora.
Paquete Tidyverse	Conjunto de paquetes diseñados para manipulación y estructuración de datos.
Posixt	La clase Date nos permitía trabajar con fechas.
as.POSIXct()	Función que convierte un tipo de dato character a clase POSIXct.
as.Posixlt()	Para crear una fecha de tipo Posixlt.
POSIXct	Subclase Posixt que almacena internamente esta cifra como un número entero.



POSIXlt	Subclase Posixt que descompone en una lista con elementos para los segundos, minutos, horas, día, mes y año.
R Studio	Es un software open source . Entorno de desarrollo integrado que nos permite utilizar con facilidad el lenguaje R.
Script	Archivo que contiene un listado secuencial de líneas de código.
Tabla de contingencia	Tabla que cuenta las observaciones por múltiples variables categóricas. Las filas y columnas de las tablas corresponden a estas variables categóricas.
Vector	Conjunto de valores (números o símbolos), todos del mismo tipo ordenados de la forma (elemento 1, elemento 2, ... , elemento n), siendo n la longitud o tamaño del vector.