

Introduccion a R y RStudio

Taller de Introduccion a R y manejo de informacion grillada

Eduardo Noriega

Noviembre de 2018



¿Qué es R?

“R es un lenguaje y entorno de libre acceso para la computación estadística y gráficos que proporciona una amplia variedad de técnicas estadísticas y gráficas: modelado lineal y no lineal, pruebas estadísticas, análisis de series temporales, clasificación, clustering, etc.”

- R core group

Descargar R y RStudio



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



<https://www.rstudio.com/products/rstudio/download/>



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

```
Rterm (64-bit)

R version 3.5.0 (2018-04-23) -- "Joy in Playing"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (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.

> -
```



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

A screenshot of the RGui (64-bit) window. The title bar says 'RGui (64-bit)'. The menu bar includes 'Archivo', 'Editar', 'Visualizar', 'Misc', 'Paquetes', 'Ventanas', and 'Ayuda'. The toolbar contains icons for file operations and running code. The 'R Console' pane shows the following text:

```
R version 3.5.0 (2018-04-23) -- "Joy in Playing"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

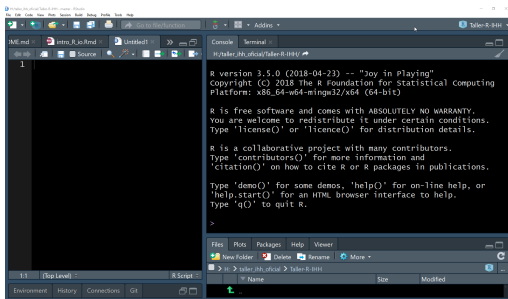
R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

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.

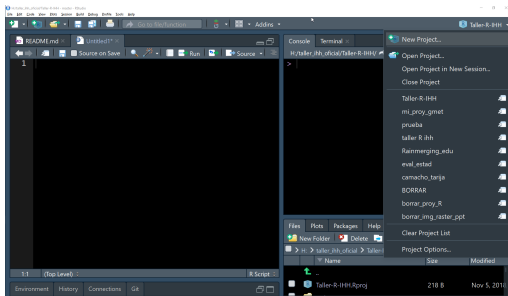
[Previously saved workspace restored]

> |
```



<https://www.rstudio.com/>

Crear proyecto



- Para poder operar de forma ordenada es recomendable crear un **Proyecto** que contenga la información de entrada, salidas, scripts, etc.
- De esa forma las rutas de acceso se vuelven relativas sin depender de un árbol de carpetas propio de cada usuario.

```
#install.packages('beep', repos = 'https://cloud.r-project.org')  
#library(beep)  
#beep(sound = 8)
```


Instalar librerías

```
install.packages('raster')  
install.packages('openxlsx')  
install.packages('ncdf4')  
install.packages('lubridate')
```

Tipos de objetos en R

Vector numérico

```
precipitacion <- c(6, 7, 24, 17, 20.5, 20.5, 20, 12, 11, 14, 38,  
                  9.5, 8, 5.5, 7)  
head(precipitacion)  
## [1] 6.0 7.0 24.0 17.0 20.5 20.5
```

Vector de caracteres

```
fecha <- c('1991-01-01', '1991-01-02', '1991-01-03', '1991-01-04',  
          '1991-01-05', '1991-01-06', '1991-01-07', '1991-01-08',  
          '1991-01-09', '1991-01-10', '1991-01-11', '1991-01-12',  
          '1991-01-13', '1991-01-14', '1991-01-15')  
head(fecha, 4)  
## [1] "1991-01-01" "1991-01-02" "1991-01-03" "1991-01-04"
```

Tipos de objetos en R

Factores

```
(valores <- factor(c('b', 'b', 'd', 'c', 'd'),  
                  levels = c('a', 'b', 'c', 'd'))  
  
## [1] b b d c d  
## Levels: a b c d  
table(valores)  
  
## valores  
## a b c d  
## 0 2 1 2
```

Variables lógicas

```
a <- 2  
b <- 3  
a == b # ¿es a igual a b?  
  
## [1] FALSE
```

Tipos de objetos en R

Matrices

```
(A <- matrix(c(-67, -16, -68.5, -17.5), byrow = FALSE, ncol = 2))  
##           [,1] [,2]  
## [1,]    -67 -68.5  
## [2,]    -16 -17.5
```

Data frames

```
datos <- data.frame(fecha, precipitacion)  
head(datos)  
##           fecha precipitacion  
## 1 1991-01-01           6.0  
## 2 1991-01-02           7.0  
## 3 1991-01-03          24.0  
## 4 1991-01-04          17.0  
## 5 1991-01-05          20.5  
## 6 1991-01-06          20.5
```

Tipos de objetos en R

Listas

```
(mi_lista <- list(matriz = A, data_frame = datos))
```

```
## $matriz
```

```
##      [,1] [,2]
```

```
## [1,]  -67 -68.5
```

```
## [2,]  -16 -17.5
```

```
##
```

```
## $data_frame
```

```
##      fecha precipitacion
```

```
## 1 1991-01-01           6.0
```

```
## 2 1991-01-02           7.0
```

```
## 3 1991-01-03          24.0
```

```
## 4 1991-01-04          17.0
```

```
## 5 1991-01-05          20.5
```

```
## 6 1991-01-06          20.5
```

```
## 7 1991-01-07          20.0
```

```
## 8 1991-01-08          12.0
```

```
## 9 1991-01-09          11.0
```

Tipos de objetos en R

Fechas

```
class(fecha)
## [1] "character"
fecha_formato <- as.Date(fecha) # ojo!
head(fecha_formato, 4)
## [1] "1991-01-01" "1991-01-02" "1991-01-03" "1991-01-04"
class(fecha_formato)
## [1] "Date"
```

Tipos de objetos en R

Funciones

Si una tarea se repite mas de dos veces es hora de crear una funcion.

```
is.leapyear <- function(year){  
  condicion_1 <- (year %% 4 == 0)  
  condicion_2 <- (year %% 100 != 0)  
  condicion_3 <- (year %% 400 == 0)  
  
  return((condicion_1 & condicion_2) | condicion_3)  
}
```

Operaciones lógicas

```
==, <, >, <=, >=, %in%
```

```
a <- 1 ; b <- c(1, 2, 3)
```

```
a == b
```

```
## [1] TRUE FALSE FALSE
```

```
a <= b
```

```
## [1] TRUE TRUE TRUE
```

```
a %in% b
```

```
## [1] TRUE
```

```
5 %in% b
```

```
## [1] FALSE
```


Indexacion

Vector

```
fecha_formato[2]
```

```
## [1] "1991-01-02"
```

Matriz y data frame

```
datos[1,2]
```

```
## [1] 6
```

```
head(datos[,2], 11)
```

```
## [1] 6.0 7.0 24.0 17.0 20.5 20.5 20.0 12.0 11.0 14.0 38.0
```

Data frame

```
datos_precip <- datos$precipitacion
```

```
head(datos_precip, 11)
```

```
## [1] 6.0 7.0 24.0 17.0 20.5 20.5 20.0 12.0 11.0 14.0 38.0
```

Listas

```
mi_lista[1]
## $matriz
##      [,1] [,2]
## [1,]  -67 -68.5
## [2,]  -16 -17.5
mi_lista[[1]]
##      [,1] [,2]
## [1,]  -67 -68.5
## [2,]  -16 -17.5
```

Información de data frames

Propiedades de los datos

```
str(datos)
```

```
## 'data.frame':    15 obs. of  2 variables:  
## $ fecha          : Factor w/ 15 levels "1991-01-01","1991-01-02",  
## $ precipitacion: num  6 7 24 17 20.5 20.5 20 12 11 14 ...
```

Algunos estadísticos

```
summary(datos$precipitacion)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	5.50	7.50	12.00	14.67	20.25	38.00

Bucles vs Vectorización

Bucles

```
salida <- vector()
for (i in 1:3) { salida[i] <- paste('indice', i, sep = '=') }
salida
## [1] "indice=1" "indice=2" "indice=3"
```

Vectorización

```
sapply(1:3, function(i) paste('indice', i, sep = '='))
## [1] "indice=1" "indice=2" "indice=3"
```

Manipulacion de informacion tabular

	A	B	C	D	E	F	G	H	I	J	K	L	M	N												
1	Estación:		San Calixto						Latitud Sud:		16° 29' 43"															
2	Departamento:		La Paz						Longitud Oeste:		68° 7' 57"															
3	Provincia:		Murillo						Altitud m/s/nm:		3658															
4																										
5	DATOS DE : PRECIPITACION DIARIA (mm)																									
6	1981																									
7	DIA	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	TOTAL												
8	1	8.0	0.3	8.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	19.2												
9	2	0.8	4.5	1.6	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.9												
10	3	11.9	0.0	3.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	4.7	21.0												
11	4	22.4	0.0	0.3	0.0	0.0	0.0	0.0	0.0	22.8	0.0	0.0	0.2	45.7												
12	5	8.1	0.0	2.1	0.0	0.0	0.0	0.0	3.7	0.0	9.4	3.2	0.0	26.5												
13	6	1.0	18.2	2.3	0.7	0.0	0.0	0.0	9.4	0.0	0.7	1.5	0.0	33.8												
14	7	0.0	29.1	13.1	3.8	0.0	0.0	0.0	1.4	0.0	5.0	0.0	0.0	52.4												
15	8	3.6	3.5	12.1	0.0	0.0	0.0	0.0	0.3	0.0	8.5	6.6	0.0	34.6												
16	9	1.1	7.4	17.8	10.2	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	37.3												
17	10	7.8	6.3	0.5	11.3	0.0	0.0	0.0	9.1	0.0	0.0	0.0	0.0	35.0												
18	11	3.8	0.0	0.0	20.5	2.2	0.0	0.0	0.0	0.6	0.0	0.0	2.0	29.1												
19	12	3.7	0.0	6.8	0.0	0.7	0.0	0.0	0.4	0.1	0.0	0.0	0.0	11.7												
20	13	5.7	19.0	0.3	1.3	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.0	27.9												
21	14	5.3	2.6	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	9.3												
22	15	1.9	3.1	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	0.0	2.2	8.1												
23	16	22.8	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.6	24.4												
24	17	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0	0.7												
25	18	9.4	0.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	11.2												
26	19	7.8	0.5	3.4	0.0	0.0	0.0	0.0	0.0	2.0	0.0	3.3	1.7	18.7												
27	20	0.5	11.5	0.0	0.0	1.3	0.0	0.0	0.0	0.7	0.0	0.0	1.1	15.1												
28	21	4.7	0.1	0.0	0.0	0.0	0.0	0.0	0.0	8.3	0.8	0.5	7.0	21.4												
29	22	0.0	5.0	0.6	0.0	0.0	0.0	0.0	0.0	0.9	3.5	0.0	0.0	10.0												
30	23	0.0	0.3	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	1.5	2.0												
31	24	0.0	13.8	0.0	0.0	0.0	0.0	0.0	0.0	8.7	6.0	0.3	2.7	31.5												
32	25	0.0	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.5	7.7	0.0												
33	26	7.1	3.2	0.3	0.0	0.0	0.0	0.0	0.0	5.5	6.6	0.0	2.0	24.7												
34	27	19.0	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.2	0.0	3.5	37.4												
35	28	2.6	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.2	0.0												
36	29	0.0	****	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	4.1												
37	30	0.0	****	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	6.0												
38	31	0.0	****	0.0	****	0.0	****	0.0	0.0	****	6.8	****	0.4	7.2												
39	SUM	159.0	137.7	75.4	48.8	4.2	0.0	0.0	27.4	51.1	70.3	28.0	42.9	644.8												
40	MED	7.2	6.9	4.2	7.0	1.4	0.0	0.0	3.0	4.6	5.0	2.3	2.5	4.4												
41	MAX	22.8	29.1	17.8	20.5	2.2	0.0	0.0	9.4	22.8	14.2	7.7	7.0	29.1												
42	MIN	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0												
43	N	22.0	20.0	18.0	7.0	3.0	0.0	0.0	9.0	11.0	14.0	12.0	17.0	133.0												
44	AÑO: 1982																									
45	DIA	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC	TOTAL												
46	1	1.0	0.0	11.7	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.5	0.0	14.2												
47	2	0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	0.0	2.1												
48	3	12.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.6												
49	4	6.4	0.0	24.2	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.6	31.5												
50	5	5.6	0.0	3.7	3.8	0.0	0.0	0.0	0.0	2.7	0.3	0.0	11.5	27.6												
51	6	2.4	0.0	0.0	11.9	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.5	15.3												
52	7	9.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.5	12.6												
53	8	6.7	1.7	0.0	0.5	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.6	10.5												
54	9	6.7	1.7	0.0	0.5	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.6	10.5												

	A	B
1	fecha	San_Calixto
2	01/01/1981	8
3	02/01/1981	0.8
4	03/01/1981	11.9
5	04/01/1981	22.4
6	05/01/1981	8.1
7	06/01/1981	1
8	07/01/1981	0
9	08/01/1981	3.6
10	09/01/1981	1.1
11	10/01/1981	7.8
12	11/01/1981	3.8
13	12/01/1981	3.7
14	13/01/1981	5.7
15	14/01/1981	5.3
16	15/01/1981	1.9
17	16/01/1981	22.8
18	17/01/1981	0
19	18/01/1981	9.4
20	19/01/1981	7.8
21	20/01/1981	0.5
22	21/01/1981	4.7
23	22/01/1981	0
24	23/01/1981	0
25	24/01/1981	0
26	25/01/1981	0
27	26/01/1981	7.1
28	27/01/1981	19
29	28/01/1981	2.6
30	29/01/1981	0
31	30/01/1981	0
32	31/01/1981	0
33	01/02/1981	0.3
34	02/02/1981	4.5
35	03/02/1981	0
36	04/02/1981	0
37	05/02/1981	0
38	06/02/1981	18.2
39	07/02/1981	29.1
40	08/02/1981	3.5
41	09/02/1981	7.4
42	10/02/1981	6.3