





# Introducción a la programación en R

Dr. José Benito Hernández Chaudary

CIMAT

Octubre/Noviembre 2019



## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
  - Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Oomo programar en R
- Gráficos con R
- 6 Resumen



# Índice

- 🚺 Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



### Nociones básicas de R

#### Nociones básicas de R

- Lo primero que nos aparece es una ventana, también llamada consola, donde podemos manejar R mediante la introducción de código.
- Sin embargo, esta no es la manera más eficiente de trabajar en R.
- A menos que estemos realizando un trabajo de mediana complejidad, será muy útil manejar todas las entradas que solicitemos a R en un entorno donde podamos corregirlas, retocarlas, repetirlas, guardarlas para continuar el trabajo en otro momento, etc.
- Esta es la función del editor de R.
- Para ello seleccionemos Nuevo script del menú Archivo y para mayor comodidad, elegimos la opción Divida horizontalmente del menú Ventanas de la consola.

### Nociones básicas de R

#### Nociones básicas de R

- La utilidad de un script o guión de trabajo radica en que podemos modificar nuestras líneas de código con comodidad y guardarlas para el futuro.
- Para ello, utilizaremos la opción Guardar o Guardar como del menú Archivo de la consola.
- Después podremos recuperar el script previamente guardado mediante la opción Abrir script del mismo menú.
- Es posible incluir comentarios que R no leerá si utilizamos líneas que comiencen con el carácter #.
- Por el contrario, si escribimos cualquier orden no antecedida de # y queremos solicitar la respuesta a R, podemos hacerlo mediante la tecla F5 situándonos en cualquier posición de esa línea (no necesariamente en el final) o la combinación de teclas Control+R.
- Asimismo, si seleccionamos con el ratón más de una línea, éstas pueden ser ejecutadas simultáneamente también con F5 o Control+R.

# Índice

- 🚺 Nociones básicas de R
  - R Studio
    - La consola de R
- 2 Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



### R Studio

#### R Studio

- RStudio es un entorno de desarrollo integrado (IDE) para R (lenguaje de programación).
- Incluye una consola, editor de sintaxis que apoya la ejecución de código, así como herramientas para el trazado, la depuración y la gestión del espacio de trabajo.
- RStudio está disponible para Windows, Mac y Linux o para navegadores conectados a RStudio Server o RStudio Server Pro (Debian / Ubuntu, RedHat / CentOS, y SUSE Linux).
- RStudio tiene la misión de proporcionar el entorno informático estadístico R. Permite un análisis y desarrollo para que cualquiera pueda analizar los datos con R.
- IDE construido exclusivo para R
  - El resaltado de sintaxis, auto completado de código y sangría inteligente.
  - Ejecutar código R directamente desde el editor de código fuente.
  - Salto rápido a las funciones definidas.



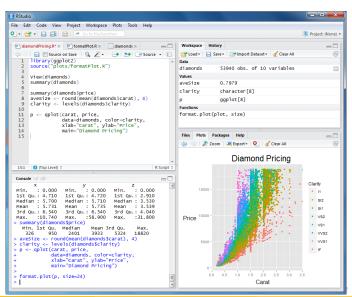
### R Studio

#### R Studio

- Colaboración
  - Documentación y soporte integrado.
  - Administración sencilla de múltiples directorios de trabajo mediante proyectos.
  - Navegación en espacios de trabajo y visor de datos.
- Potente autoría y depuración
  - Depurador interactivo para diagnosticar y corregir los errores rápidamente.
  - Herramientas de desarrollo extensas.
  - Autoría con Sweave y R Markdown.



### R Studio



# Índice

- 🚺 Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



### Sobre la consola podemos ejecutar diferentes acciones

- Ayuda
- Ver las librerías y objetos disponibles



### Manera clásica de consultar la ayuda

```
help(sd)
```

?sd

help.search("solve") # Archivos relacionados con "solve"

help() # Avuda para caracteres especiales o palabras reservadas



### Librerías y objetos disponibles

library() # Muestra las librerías disponibles que pueden ser cargadas search() # Muestra las librerías ya cargadas y disponibles ls() # Objetos usados





### Se puede usar R como una simple calculadora

```
5-1+10 # suma y resta
7*10/2 # multiplica y divide
pi # el número pi
sqrt(2) # Raíz cuadrada
```



### Crear objetos mediante asignaciones

```
x <- 5 # El objeto x toma el valor 5
x # imprime x
6 = x \# x ahora toma el valor 6
(x <- pi) # asigna el número pi a x e imprime
```





### R es sensible a mayúsculas y minúsculas

- # Dos objetos diferentes
- b < -2
- B <-4

#### Los objetos pueden ser de tipo:

- Numéricos, carácter, lógicos, factores
- Vectores, matrices, listas.
- Funciones, entre otros.



### Comandos útiles

```
ls() # Muestra los objetos usados
rm(b) # Borra objetos usados
ls()
```

R tiene un gran número de funciones. Al crear objetos (como x) hay que evitar los nombres de funciones. Para descubrir si un nombre es de una función predeterminada de R, se pone simplemente el nombre que queremos usar para el objeto:





# Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



## Tipos de objetos en R

#### Tipos de objetos en R

La información manipulada en R es en forma de objetos. Ejemplos de objetos son vectores de valores numéricos (reales) o valores complejos, vectores de valores lógicos y vectores de caracteres. Las mismas funciones del R son objetos. R también opera con objetos llamados listas, que son del modo lista. Estos son secuencias ordenadas de objetos que individualmente pueden ser de cualquier tipo. Concretamente, vamos a hablar de:

- Vectores.
- Matrices.
- Listas y Factores.
- Hojas de datos (Data frame).

# Índice

- Nociones básicas de F
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- Resumen



#### **Vectores**

- Los elementos de un vector pueden ser: enteros, decimales, caracteres, etc.
- R efectúa las operaciones aritméticas entre vectores componente a componente: si sumamos dos vectores de igual longitud el resultado es otro vector de la misma longitud, cuyas componentes son la suma de las componentes de los vectores que sumamos.
- De manera similar ocurre si realizamos cualquier otra operación aritmética, incluyendo la potenciación.
- Veamos algunos ejemplos.

# **Vectores**

```
x<-0:19 # una secuencia de 20 números
x[5] # muestra el elemento en la 5ta. posición
x<-seq(1, 10, length = 5)
seq(1,20,by=2) # Secuencia de dos en dos
# Vector de ceros.
numeric(6)
integer(3)
rep(2,6) # Vector de constantes.</pre>
```

```
Vector a través de la función "c"
# Vector de números reales y lo asignamos a x
x < -c(1.3, 2.5, 3.6, 4.7, 8.2)
v < -c(x, 0, x) # Generamos un vector a partir de otro vector.
# Se pueden agregar nombres a los vectores.
z < -c(1.3.45)
names(z)
names(z) <-c("primero", "segundo", "tercero")</pre>
z
```

# Vector lógico

```
x < -1:5
cond1 < -x < 4
cond1
cond2 < -x > = 3
cond1 & cond2 # Hacemos una 'y' lógica
logical(3) # Vector de tres elementos lógicos
```

#### **Vectores**

Así como generamos los vectores lógicos también podemos generar caracteres

### Generando caracteres

```
(x<-character(3)) # vector de 3 entradas vacías
# Vector con las 4 primeras letras del alfabeto
x2=c("A","B","C","D")
# Otra forma
x3=LETTERS[1:4]</pre>
```

#### **Vectores**

Para saber cuales son los atributos de una variable u objeto podemos utilizar las siguientes funciones:

mode(), is.logical(x) ; is.numeric(x) ; is.integer(x), etc.

Algunos valores que pueden aparecer cuando trabajamos con vectores u otros objetos son NA, NaN, Inf. Estas tres cantidades tienden a confundirse. Vamos a tratar de aclarar su significado mediante un ejemplo.

```
Ejemplo

(x<-c(NA, 0 / 0, Inf - Inf, Inf, 5))

NA: # representa un elemento ausente

NAN: # Son las indeterminaciones

Inf: # representa al infinito
```

#### **Vectores**

- Podemos conocer la longitud o dimensión de un vector utilizando el comando length ()
- Por ejemplo, podemos ver la longitud de uno de los vectores que ya hemos creado length (y)
- Podemos convertir los vectores de numéricos a lógicos, de lógicos a enteros y viceversa.
   Esto se hace con las funciones as.numeric, as.logical, as.integer, etc.



```
h1=c(1.3,0.6,-3,8);h1
h2=as.integer(h1);h2
as.logical(h2)
```



#### Operaciones con vectores

Vamos a realizar ahora algunas operaciones con vectores tales como sumas, productos, potencias, etc.

# **Æ**Ejemplo

```
a<-5:2
b<-(1:4)*2
a
[1] 5 4 3 2
b
[1] 2 4 6 8
```

```
Ejemplo
 [1] 7 8 9 10
 a - b
 [1] 3 0 -3 -6
 a * b
 [1] 10 16 18 16
 a / b
 [1] 2.50 1.00 0.50 0.25
 a \Lambda b
 [1] 25 256 729 256
```

#### Operaciones con vectores

- Es posible realizar operaciones aritméticas entre un vector y un escalar
- También es posible realizar operaciones con más de dos vectores simultáneamente

```
Ejemplo

# Producto de un escalar por un vector
2 * a
[1] 10 8 6 4
# Operación con más de dos vectores
d<-rep(2,4)
a + b + d
[1] 9 10 11 12
```

#### Ejemplo. Operaciones con vectores

Veamos ahora un ejemplo más interesante. Supongamos que queremos evaluar la función

$$f(x,y) = \log\left(\frac{x^2 + 2y}{(x+y)^2}\right)$$

para varios valores de x e y. Una posibilidad es tomar cada par de valores y calcular f(x,y). Podemos, sin embargo, aprovechar la forma en la cual trabaja R con vectores para realizar todas las operaciones con una sola evaluación. Comenzamos por crear los vectores que contienen los valores de interés para x e y.

```
Ejemplo
 # Definimos los vectores x e y
 (x<-10:6)
 [1] 10 9 8 7 6
 (v < -seq(1, 9, 2))
 [1] 1 3 5 7 9
 # Definimos f(x,y) en términos de estos vectores.
 # Guardamos los resultados en z
 (z < -\log(x \wedge 2 + 2 \star y) / (x + y) \wedge 2)
 [1] -0.1708177 -0.5039052 -0.8258336
 [4] -1.1349799 -1.4271164
```

#### Operaciones con vectores

Funciones de uso común para vectores:

| Nombre              | Operación                                   |
|---------------------|---|
| length()            | longitud                                    |
| sum()               | suma de las componentes del vector          |
| prod                | producto de las componentes del vector      |
| cumsum(), cumprod() | suma y producto acumulados                  |
| max(), min()        | máximo y mínimo del vector                  |
| cummax(), cummin()  | máximo y mínimo acumulados                  |
| sort()              | ordena el vector                            |
| diff()              | calcula la diferencia entre las componentes |
| which(x==a)         | vector de los índices de x para los         |
|                     | cuales la comparación es cierta             |

Table: Funciones vectoriales

#### Operaciones con vectores

| Nombre       | Operación  |
|--------------|--|
| which.max(x) | índice del mayor elemento                        |
| which.min(x) | índice del menor elemento                        |
| range(x)     | valores del mínimo y el máximo de x              |
| mean(x)      | promedio de los elementos de x                   |
| media(x)     | mediana de los elementos de x                    |
| round(x,n)   | redondea los elementos de x a n decimales        |
| rank(x)      | rango de los elementos de x                      |
| unique(x)    | vector con las componentes de x sin repeticiones |

Table: Funciones vectoriales

# Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



#### **Matrices**

- Una matriz es un arreglo rectangular de celdas, cada una de las cuales contiene un valor.
- Para crear una matriz en R es posible usar la función matrix, cuya sintaxis es
  matrix (datos, nrow, ncol, byrow=F, dimnames = NULL) donde nrow y ncol
  representan, respectivamente, el número de filas y columnas de la matriz.
- Sólo el primer argumento es indispensable.
- Si no aparecen ni el segundo ni el tercero, los datos se colocan en una matriz unidimensional, es decir, en un vector.
- Si sólo uno de los valores se incluye, el otro se determina por división y se asigna.



Por ejemplo, para definir la matriz

$$\left(\begin{array}{ccc}
1 & 4 & 7 \\
2 & 5 & 8 \\
3 & 6 & 9
\end{array}\right)$$

Usaremos

matriz<-matrix(c(1,2,3,4,5,6,7,8,9),nrow=3,ncol=3)

#### **Matrices**

- Las dimensiones de la matriz pueden obtenerse mediante la función dim ( ).
- Si queremos llegar a elementos concretos de una matriz lo haremos utilizando corchetes para indicar las filas y columnas. Por ejemplo,

```
matriz[2,3] # devolvería el valor 8,
matriz[1:2,2:3]
matriz[,c(1,3)] # devolverá ??
```

Tanto para vectores como para matrices, funcionan las operaciones suma y diferencia, si
queremos multiplicar dos matrices usaremos matriz\*\*\*matriz mientras que
matriz\*matriz devuelve la multiplicación elemento por elemento.

```
Matrices
matrix(1:6)
    [, 1]
[1,] 1
[3,] 3
[5, ] 5
[6,] 6
matrix(1:6, nrow=3)
    [, 1] [, 2]
[2,] 2 5
[3,] 3 6
```

#### **Arreglos**

- Un arreglo (array) de datos es un objeto que puede ser concebido como una matriz multidimensional (hasta 8 dimensiones).
- Una ventaja de este tipo de objeto es que sigue las reglas que hemos descrito para las matrices.
- La sintaxis para definir un arreglo es

array(data, dim).

- Las componentes data y dim deben presentarse como una sola expresión.
- Por ejemplo c (2, 4, 6, 8, 10) o c (2, 3, 2):
   x <- array (1:24, c(3, 4, 2))</li>
   produce un arreglo tridimensional: la primera dimensión tiene tres niveles, la segunda tiene cuatro y la tercera tiene dos.
- Al imprimir el arreglo R comienza con la dimensión mayor y va bajando hacia la dimensión menor, imprimiendo matrices bidimensionales en cada etapa.





### Arreglos (Vectores o matrices).

```
array(1:4,6) # Vector de tamaño 6
array(1:6,c(2,5)) # Matriz 2x5, se llena por columnas
# Matrices (Por defecto llena por columnas)
matrix(c(4,5,9,52,67,48),nrow=2,ncol=3)
# Para llenarla por filas, se le agrega
matrix(c(4.5.9.52.67.48),nrow=2,ncol=3.bvrow=TRUE)
# Colocando vectores como columna
(y<-cbind(letters[1:4], LETTERS[1:4]))
```

## Índice

- Nociones básicas de F
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



#### Listas

En R también podemos crear listas y factores. En R, una lista es un objeto consistente en una colección ordenada de objetos, conocidos como componentes. No es necesario que los componentes sean del mismo modo ni tampoco que tengan la misma estructura, así una lista puede estar compuesta de, por ejemplo, un vector numérico, un valor lógico, una matriz y una función.



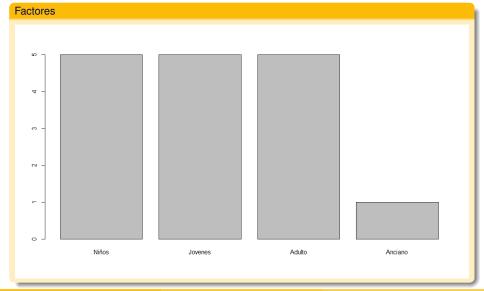
```
lista<-list(Marca="Chevrolet", Modelo="Aveo", n.puertas=5,
           Año=c(2006,2007))
lista
# Seleccionamos posiciones de la lista
lista[[1]] # Posición 1 de la lista
lista[[4]][2] # Posición 4 de la lista, subposición2
# También podemos referirnos a las posiciones por los nombres.
lista$Marca
lista$Modelo
```

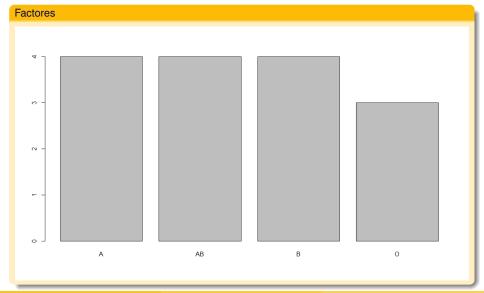
#### **Factores**

- Los factores son un tipo especial de vectores que permiten analizar un conjunto de datos clasificados según las características que definen el factor.
- Los factores, son vectores unidimensional que incluye los valores correspondientes a una variable categórica, pero también los diferentes niveles posibles de esta variable.
- Los factores permiten trabajar con modelos y gráficas de variables categóricas.
- Por eso son importantes.
- Hay dos formas de hacer factores, la primera es usar la función g1 (), en donde especificamos el número de niveles, el número de repeticiones y el número total de datos respectivamente, y con el argumento labels se especifica el nombre de los niveles.
- La segunda forma es utilizar es usando la función factor (), en donde especificamos los datos tipo caracter, el número de niveles y el número total de datos.

# **Factores**

```
# Con la funcion gl()
edades <- ql (4,5,16, labels = c ("Niños", "Jovenes", "Adulto", "Anciano"))
edades
# Grafica del factor
plot (edades)
# Con la funcion factor()
sangre<-factor(rep(c("A", "B", "AB", "O"), 4, 15))</pre>
sangre
# Grafica del factor
plot (sangre)
```





#### **Factores**

En ocasiones es importante ordenar por nivel de importancia a los factores a estos se le llaman factores ordenados, para eso ocupamos la función ordered (): en donde colocamos primero el vector que queremos ordenar, seguido del argumentos levels en donde colocaremos primeros los niveles de mayor importancia.

## Factores

```
escolaridad<-factor(rep(c("MedioSuperior", "Primaria", "Secundaria", "Superior", "Prescolar"),5,15))
escolaridad
ordered(escolaridad,levels=c("Superior", "MedioSuperior", "Secundaria", "Primaria", "Prescolar"))
```

## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



### Data frame

#### Data frame

Una hoja de datos (Data frame) es una lista que pertenece a la clase "data.frame". Hay restricciones en las listas que pueden pertenecer a esta clase, en particular:

- Los componentes deben ser vectores (numéricos, cadenas de caracteres, o lógicos), factores, matrices numéricas, listas u otras hojas de datos.
- Las matrices, listas, y hojas de datos contribuyen a la nueva hoja de datos con tantas variables como columnas, elementos o variables posean, respectivamente.
- Los vectores numéricos y los factores se incluyen sin modificar, los vectores no numéricos se fuerzan a factores cuyos niveles son los únicos valores que aparecen en el vector.
- Los vectores que constituyen la hoja de datos deben tener todos la misma longitud, y las matrices deben tener el mismo tamaño de filas

Las hojas de datos pueden interpretarse, en muchos sentidos, como matrices cuyas columnas pueden tener diferentes modos y atributos. Pueden imprimirse en forma matricial y se pueden extraer sus filas o columnas mediante la indexación de matrices.



Edo.Civil=c("C", "S", "S", "C"))

### Data frame

dataf



Por defecto, tanto las listas como los data frame están invisibles en R, para hacerlas visibles y dejar de usar \$ usamos la función attach y para hacerlas invisibles nuevamente usamos detach.



attach(lista) Marca attach(dataf) Edad

### Data frame



- Vamos a ver cómo se construye una hoja de datos con los datos de 3 personas, que incluye el color de sus ojos como factor, su peso y su altura.
- Empezaríamos definiendo el color de los ojos:

- Supongamos que los pesos y las alturas son, respectivamente: 68, 75, 88 y 1.65, 1.79, 1.85.
- Entonces, definiríamos la hoja de datos mediante:

### Data frame

#### Data frame

- Podemos forzar a que una matriz se convierta en una hoja de datos mediante: datos2<-as.data.frame(matriz)</li>
- Si ponemos names(datos2) veremos los nombres que para las variables ha elegido por defecto R:

```
[1] "V1" "V2" "V3"
```

 Si queremos modificar esos nombres de las variables, podemos usar de nuevo la función names ():

```
names(datos2)<-c("Variable 1","Variable 2","Variable 3")</pre>
```

 Podemos usar el operador \$, de la siguiente manera. Para obtener los datos de la variable Color.ojos, por ejemplo, escribiríamos datos\$Color.ojos

## Tipos de objetos en R

#### **Atributos**

- Un factor es una variable categórica.
- Una matriz es una tabla con dos dimensiones.
- Un arreglo es una tabla con k dimensiones.
- Para una matriz o un arreglo los elementos deben ser todos del mismo tipo.
- Una hoja de datos (data frame) es una tabla formada por vectores y/o factores de la misma longitud pero posiblemente de modos distintos.
- Un ts es un conjunto de datos correspondiente a una serie temporal y tiene atributos adicionales como frecuencia y fechas.
- Finalmente, uno de los objetos más útiles es la lista (list), que puede ser usada para combinar cualquier colección de objetos de datos en un solo objeto (incluyendo otras listas).



## Tipos de objetos en R

#### Tipos de objetos en R

| Objeto        | Modos   | Varios modos |  |
|---------------|---|--------------|--|
| Vector        | Numérico, caracter, complejo o lógico           | No           |  |
| Factor        | Numérico o caracter (categórico)                | No           |  |
| Matriz        | Numérico, caracter, complejo o lógico           | No           |  |
| Arreglo       | Numérico, caracter, complejo o lógico           | No           |  |
| Hoja de datos | Numérico, caracter, complejo o lógico           | Si           |  |
| ts            | Numérico, caracter, complejo o lógico           | No           |  |
| Lista         | Numérico, caracter, complejo o lógico, función, | Si           |  |



## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



### Lectura/Escritura de datos

#### Lectura/Escritura de datos

- R utiliza el directorio de trabajo para leer y escribir archivos.
- Para saber cual es este directorio puede utilizar el comando getwd () (get working directory).
- Para cambiar el directorio de trabajo, se utiliza la función setwd ().
- Por ejemplo, setwd("C:/data") o setwd("/Clases/Maestria INEGI").
- Es necesario proporcionar la direccion ('path') completa del archivo si este no se encuentra en el directorio de trabajo.

## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



#### La función read.table()

Para poder leer una hoja de datos directamente, el archivo externo debe reunir las condiciones adecuadas. La forma más sencilla es:

- La primera linea del archivo debe contener el nombre de cada variable de la hoja de datos.
- En cada una de las siguientes líneas, el primer elemento es la etiqueta de la fila, y a continuación deben aparecer los valores de cada variable.

Si el archivo tiene un elemento menos en la primera línea que en las restantes, obligatoriamente será el diseño anterior el que se utilice. A continuación aparece un ejemplo de las primeras líneas de un archivo, **datos.casas**, con datos de viviendas, preparado para su lectura con esta función.

| Uso | de la fu | ınción reac | d.table | <b>e</b> () |      |       |
|-----|----------|-------------|---------|-------------|------|-------|
|     | Precio   | Superficie  |         |             | Años | Calef |
| 01  | 52.00    | 111.0       | 830     | 5           | 6.2  | no    |
| 02  | 54.75    | 128.0       | 710     | 5           | 7.5  | no    |
| 03  | 57.50    | 101.0       | 1000    | 5           | 4.2  | no    |
| 04  | 57.50    | 131.0       | 690     | 6           | 8.8  | no    |
| 05  | 59.75    | 93.0        | 900     | 5           | 1.9  | no    |
|     |          |             |         |             |      |       |

#### Lectura de datos: Leer un archivo

Vamos a leer una tabla guardada en un archivo de texto usando la función read.table(). El archivo se llama MEXpob.txt que tiene una tabla que muestra la población (en millones) de México desde 1985.



### La función read.table()

MEXpob<-read.table("MEXpob.txt", header=TRUE)
MEXpob
plot(Pob. ~ Año, data=MEXpob, pch=16)</pre>

#### La función scan ()

La función scan es más flexible que read.table. A diferencia de esta última es posible especificar el modo de las variables:

## 🚁 La función scan()

misdatos<-scan("entrada.txt", list(" ",0,0))
misdatos</pre>

#### La función read. fwf()

La función read.fwf () puede usarse para leer datos en archivos en formato fijo ancho:

## La función read.fwf()

misdatos2 <- read.fwf("datos.txt", widths=c(1, 4, 3)) misdatos2</pre>

## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- 6 Resumen



#### Guardar datos

- Podemos salvar los objetos que hemos creado hasta ahora de la siguiente manera:
  - Archivo → Guardar área de trabajo
  - Colocamos el nombre .Rdata, ejemplo: clase1.RData.
- También podemos salvar todos los comandos ejecutados:
  - Archivo → Guardar histórico
  - Colocamos el nombre .Rdata, ejemplo: claseh1
- Para abrir el archivo guardado puedes dar doble click sobre el archivo y una vez abierto se realiza lo siguiente:
  - ◆ Archivo → Cargar área de trabajo. Clase1.RData.
- Para abrir el histórico:
  - Archivo → Histórico. claseh1



#### Guardar datos

- La función write.table guarda el contenido de un objeto en un archivo.
- El objeto es típicamente una hoja de datos ('data.frame'), pero puede ser cualquier otro tipo de objeto (vector, matriz,...).
- Una manera sencilla de escribir los contenidos de un objeto en un archivo es utilizando el comando write (x, file="data.txt"), donde x es el nombre del objeto (que puede ser un vector, una matrix, o un arreglo).
- Esta función tiene dos opciones: nc (o ncol) que define el número de columnas en el archivo (por defecto nc=1 si x es de tipo caracter, nc=5 para otros tipos), y append (logico) que agrega los datos al archivo sin borrar datos ya existentes (TRUE) o borra cualquier dato que existe en el archivo (FALSE, por defecto).
- Para guardar un grupo de objetos de cualquier tipo se puede usar el comando save (x, y, z, file= "xyz.RData").

#### Guardar datos

- Para facilitar la transferencia de datos entre diferentes máquinas se pueden utilizar la opción ascii = TRUE.
- Los datos (denominados ahora como un workspace o "espacio de trabajo" en terminología de R) se pueden cargar en memoria más tarde con el comando load ("xyz.RData").
- La función save.image() es una manera corta del comando save(list=ls(all=TRUE), file=".RData") (guarda todos los objetos en memoria en el archivo.RData).

#### Guardar datos

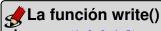
Vamos a guardar las tablas creadas anteriormente con la función read.table().

### La función write.table()

write.table(PreciosCasas, "./Datos/PreciosCasas2.txt") write.table(MEXpob, "./Datos/MEXpob2.txt")

#### Guardar datos

Aparte de utilizar la función write.table(), podemos utilizar la función write(). Supongamos que queremos guardar en un fichero el contenido de un vector



x <- c(1,2,3,4,5)
write(x, "./Datos/x.txt")</pre>

El contenido del fichero x.txt será el vector 1 2 3 4 5

#### Guardar datos

Si lo que queremos es escribir una matriz con este procedimiento, hay que ser más cuidadosos. Creamos una matriz para probar y la escribimos en el fichero xx.txt:

## La función write()

```
x <- matrix(1:9, ncol = 3, byrow = T)
x
write(t(x), "./Datos/xx.txt", ncol = ncol(x))</pre>
```

## Lectura/Escritura de datos

#### Lectura/Escritura de datos

| Función       | Descripción   |
|---------------|---|
| data.frame()  | Lee datos de un archivo de datos con formato específico             |
| scan()        | Lee datos de un archivo de datos                                    |
| write.table() | Escribe datos en un fichero con formato de tabla de datos           |
| write()       | Escribe datos en un fichero .txt                                    |
| write.csv()   | Escribe datos en un fichero .csv (excel)                            |
| save()        | Guarda un grupo de objetos en un archivo .RData                     |
| save.image()  | Guarda todos los objetos en memoria del espacio de trabajo (.RData) |
| load()        | Carga en memoria archivos .RData                                    |

## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- Resumen



#### Como programar

Al igual que otros lenguajes de programación, R incluye la ejecución de condicionales, bucles y esas construcciones. En esta parte se discute brevemente estas construcciones. Debido a su rica colección de funciones y paquetes y por su enfoque orientado a objetos, se evitará la programación en R tanto como sea posible. Hay, sin embargo, situaciones en las que tendremos que confiar en la programación.

- Para programar usaremos un Script. Entramos en:
   Archivo → Nuevo Script
- Para guardar realizamos lo siguiente:
   Archivo → Guardar como clase1



### **Condicional:**

Un condicional se ejecuta de la siguiente manera:

```
if (condición) {
   ejecutar algo
} else {
     ejecutar algo si la condición anterior
     no se verifica
}
```



```
z=1:10
# Un ejemplo de if
z1=sample(z,1)
if (z1>7) {w=1} else {w=0}
# Un ejemplo con condición lógica
x <- TRUE
if(x) y <- 1 else y <- 0
y</pre>
```



### **Condicional:**

Si se tiene más de una instrucción a ejecutar es necesario usar las llaves

```
# Dos asignaciones a ejecutar
x <- FALSE
if(!x){y <- 1 ; z <- 2}
# Dos condiciones a verificar
# Raíz n-ésima de un número real
n=5; x=-32
if(n%%2 == 1 || x >=0){
    sign(x)*abs(x) \(1/n)
} else{
    NaN
}
```

#### Ciclos:

Para repetir una acción se pueden usar bucles (ciclos) usando: for, repeat y while. Se puede salir de un ciclo en cualquier instante mediante el uso de break

- Usando un for for (condición) {acción a ejecutar}
- El while se ejecuta de la siguiente manera while (condición) {Acción a ejecutar}

```
# Usando un 'for' y creando vectores lógicos
x1 <- as.logical(as.integer(runif(5, 0, 2)));
x1
y1 <- vector(); y1
for(i in 1 : length(x1)){
   if(x1[i]){y1[i] <- 1}
   else {y1[i] <- 0}}
y1</pre>
```

```
# Un ejemplo de 'repeat' y 'break'
x2=1:10; x3=10:20; i=1
repeat{
    y=x2[i]+x3[i]
    i=i+1
    if(i==10) break
}
y
```

#### **Funciones:**

- En R también se pueden crear funciones, estas pueden tener argumentos o no, un cuerpo y en general arroja valores.
- El uso de una función en R es similar al uso matemático, en matemáticas escribimos y = f(x) y en R: y<-f(x)</li>
- La forma general de una función es

```
nombre.funcion <- function(argumentos)
{
    acciones a ejecutar y valor a retornar</pre>
```

- El valor a retornar puede ser cualquier objeto: vector, valor numérico, un gráfico etc.
- Los valores pueden ser retornados colocando simplemente el nombre de la variable que contiene la información o usando la función print



Vamos a definir una función llamada cubo que toma un número y lo eleva a la potencia tres.

```
# Definimos la función cubo
cubo <- function(x)
{return(x\lambda3)}
# Ejecutamos la función cubo con x=2
cubo(2)
[1] 8</pre>
```

#### **Funciones**

- Una función puede tener muchos argumentos y en esos casos es cómodo no tener que incluir todos los parámetros.
- En esta situación es conveniente poder asignar valores 'por defecto' o automáticos a los parámetros. Esto es posible en R usando el signo =.
- Si llamamos una función sin especificar explícitamente el parámetro, se usará su valor 'por defecto'.

```
g Ejemplo
```

```
# Definimos la función ff

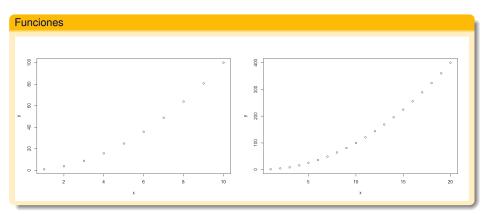
ff <- function (x = 1:10, y = (1:10)\wedge2, showgraph = T) {
    if (showgraph) plot (x,y)
    else print (cbind (x,y))
    return (invisible ())
}

# Ejecutamos la función ff

ff (1:10, (1:10)\wedge2, T)

ff (1:10, (1:10)\wedge2)

ff (1:20, (1:20)\wedge2)
```



## Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con R
- Resumen



- Una de las mayores ventajas de R es su capacidad gráfica.
- Para ver una demostración usamos los siguientes comandos:

```
demo(graphics)
demo(persp)
demo(image)
```

- Posiblemente es la función gráfica más usada en R es plot.
- Su efecto depende de la sintaxis y el argumento que usemos.
- A continuación veremos varios ejemplos de gráficos con R.

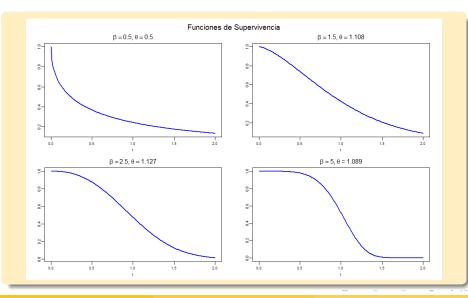
# \$

#### Ejemplo. Distribucion Weibull

# \$

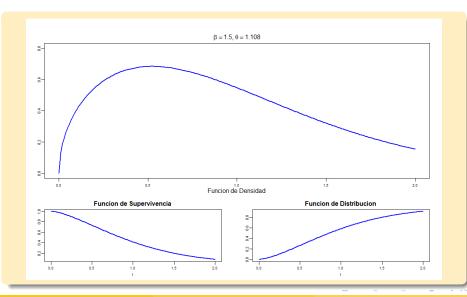
#### **Ejemplo. Distribucion Weibull**

```
b3 <- 2.5; t3 <- 1/gamma(1+1/b3)
plot(tt, exp(-(tt/t3) \hbab), xlim=c(0,2), cex.axis=.7, cex.lab=.7,
    mgp=c(1.5,.5,0), lwd=2, col="blue", type="1", xlab="t",
    ylab="",
    main= expression(paste(beta==2.5, ", ", theta==1.127)))
b4 <- 5; t4 <- 1/gamma(1+1/b4)
plot(tt, exp(-(tt/t4) \hbab), xlim=c(0,2), cex.axis=.7, cex.lab=.7,
    mgp=c(1.5,.5,0), lwd=2, col="blue", type="1", xlab="t",
    ylab="",
    main=
        substitute(paste(beta == 5, ", ", theta, " = ",t4),
        list(t4=round(t4,3)))
mtext("Funciones de Supervivencia", outer=TRUE, cex=1.2)
# (en esta grafica, notar en la ultima, el uso de substitute)</pre>
```

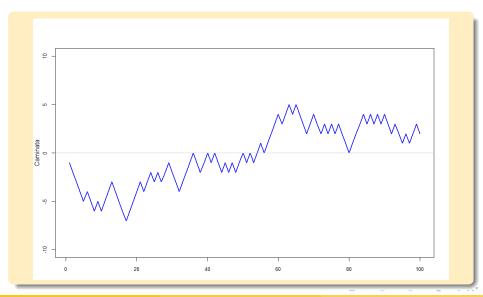




#### Ejemplo. Uso de la función layout()



# 🚀 Ejemplo. Caminata aleatoria

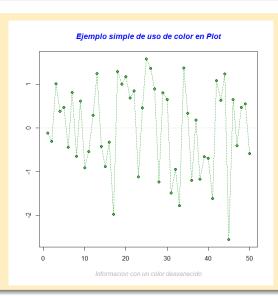


#### Gráficos con R

Los siguientes gráficos fueron tomados del demo (graphics)

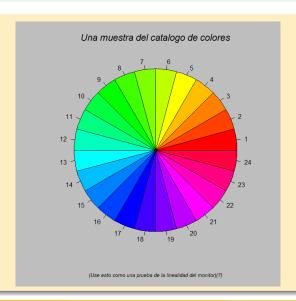
## \$

#### Ejemplo. Una grafica simple

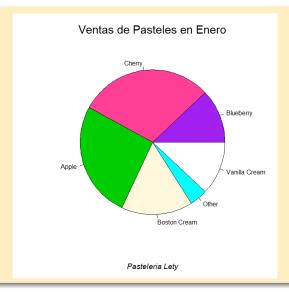


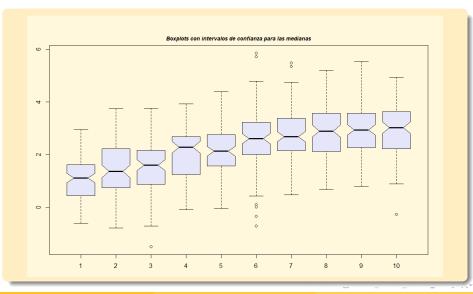


#### Ejemplo. Diagrama de pastel.



# Ejemplo. Diagrama de pastel (de nuevo).

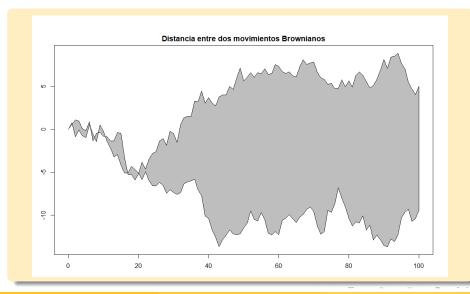






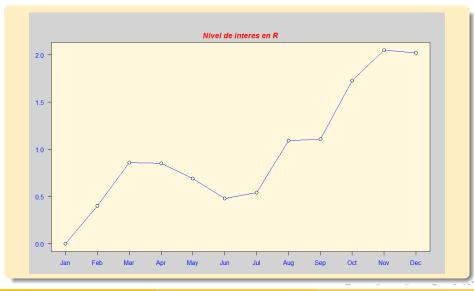
#### Ejemplo. Área sombreada entre dos gráficas.

```
par(bg="white")
n <- 100
x <- c(0,cumsum(rnorm(n)))
y <- c(0,cumsum(rnorm(n)))
xx <- c(0:n, n:0)
yy <- c(x, rev(y))
plot(xx, yy, type="n", xlab="Tiempo", ylab="Distancia")
polygon(xx, yy, col="gray")
title("Distancia entre dos movimientos Brownianos")</pre>
```





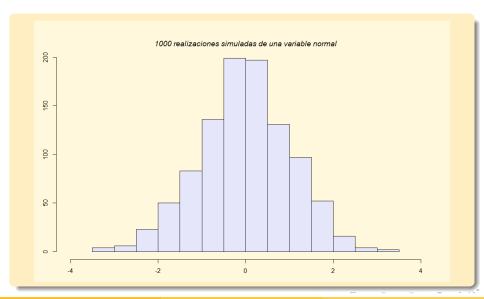
#### Ejemplo. Gráficas tipo Excel, o algo parecido.





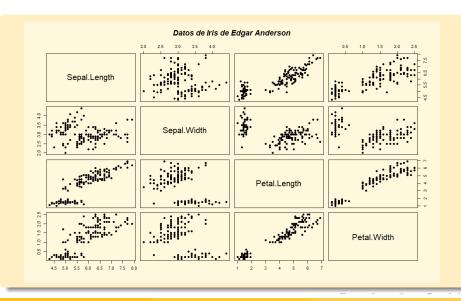
#### Ejemplo. Histograma.

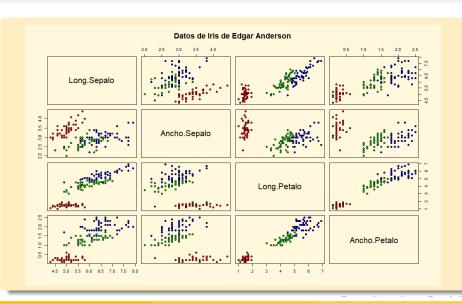
```
par(bg="cornsilk")
x <- rnorm(1000)
hist(x, xlim=range(-4, 4, x), col="lavender", main="",
    ylab="Frecuencia")
title(main="1000 realizaciones simuladas de una variable normal",
    font.main=3)</pre>
```



```
# Unicolor
pairs(iris[1:4], main="Datos de Iris de Edgar Anderson",
font.main=4, pch=19)
#--
# Colores diferentes para cada especie.
aa <- iris
names(aa) <- c("Long.Sepalo", "Ancho.Sepalo",
"Long.Petalo", "Ancho.Petalo",
"Especie")
pairs(aa[1:4], main="Datos de Iris de Edgar Anderson", pch=21,
```

bg = c("red", "green3", "blue") [unclass(iris\$Species)])





# # volcano es la matriz 87 x 61 de elevaciones del volcán # Maunga Whau en NZ x <- 10\*1:nrow(volcano) y <- 10\*1:ncol(volcano) lev <- pretty(range(volcano), 10) par(bg = "lightcyan") pin <- par("pin") xdelta <- diff(range(x)) ydelta <- diff(range(y)) xscale <- pin[1]/xdelta yscale <- min(xscale, yscale)

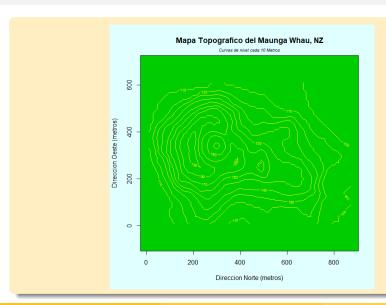
xadd <- 0.5\*(pin[1]/scale - xdelta)
yadd <- 0.5\*(pin[2]/scale - ydelta)
plot(numeric(0), numeric(0),</pre>

type = "n", ann = FALSE)

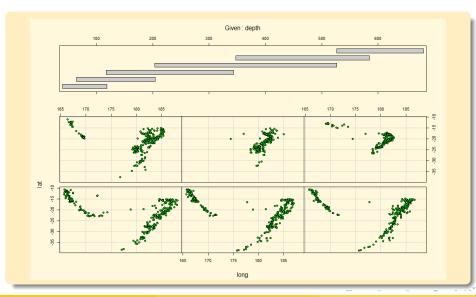
xlim = range(x) + c(-1,1) \* xadd, ylim = range(y) + c(-1,1) \* yadd,



#### Ejemplo. Gráfica de contornos



```
# El conjunto de datos quakes es un data frame con 1000
# observaciones en 5 variables:
# lat = Latitud del evento
# long = Longitud
# depth = Profundidad (km)
# mag = Magnitud en escala de Richter
# stations = Numero de estaciones reportando el evento
par(bg="cornsilk")
coplot(lat ~ long|depth, data = quakes, pch = 21, bg = "green3")
```



#### Gráficos con R

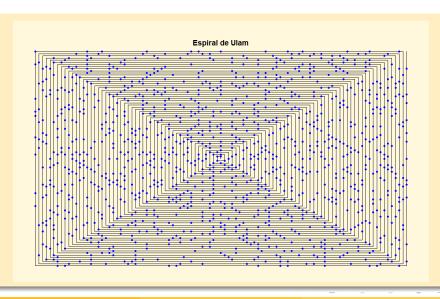
A continuación presentamos algunas figuras interesantes, cuyo código en R es un poco más elaborado, aunque las funciones gráficas son los que ya hemos usado.



```
Ejemplo. Espiral de Ulam.
# prim = Un programa para calcular los primeros n primos
prim <- function(n){
   if (n==1) {return(2)}
   primos <- 2
   notvet <- TRUE
   probar <- 2
   while (notyet) {
     probar <- probar + 1
     pritst <- primos[ primos<=sqrt(probar) ]</pre>
     aa <- (probar %% pritst)</pre>
     if (anv(aa==0)) (next)
     primos <- c(primos, probar)</pre>
     if ( length(primos) == floor(n) ) { return(primos) }
```

# **Ejemplo. Espiral de Ulam.**

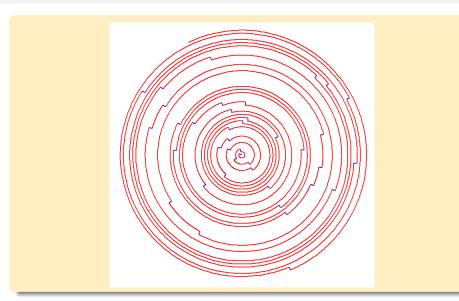
```
m < -100
pp \leftarrow prim((m+1) \land 2)
ii < - seq(3, m+1, by=2)
jj <- length(ii)</pre>
par(mar=c(0,0,0,0)+1); xylim <- c(1,m+1)
plot(1,1,xlim=xylim, ylim=xylim, type="n", xaxt="n", yaxt="n",
     bty="n", xlab="", vlab="")
aa <- c(floor(m/2)+1, floor(m/2)+1)
for(k in 1:jj){
   r \leftarrow ii[k]
   co \leftarrow cbind(c(rep(r,r), (r-1):2, rep(1,r), 2: (r-1)),
       c(r:1, rep(1, r-2), 1:r, rep(r, r-2)))
   co <- co + (ii-k)
   n \leftarrow dim(co)[1]
   uu \leftarrow (r \land 2) : ((r-2) \land 2)
   rr \leftarrow is.element(uu[-(n+1)],pp)
   bb <- co[n,]
   segments(aa[1], aa[2], bb[1], bb[2], col="black", lwd=1)
   aa <- co[1,]
```



## \$

#### Ejemplo. Laberinto circular

```
M \leftarrow 40; m \leftarrow 120; n \leftarrow M; xylim \leftarrow .95*c(-M, M)
par(mar=c(0,0,0,0)+.6)
plot(0,0, type="n", xlim=xylim, ylim=xylim, xaxt="n", yaxt="n",
    xlab="", ylab="", btv="n")
pp < -c(0,0)
tet1 <- runif(1,min=0,max=2*pi)
for( r in 1:n ) {
   gg <- r*c(cos(tet1), sin(tet1))</pre>
   segments(pp[1],pp[2],qq[1],qq[2], col="blue",lwd=2)
   tet2 <- tet1 + runif(1, min=0, max=2*pi)
   ts <- seg(tet1, tet2, length=200)
   nc <- r*cbind( cos(ts), sin(ts) )</pre>
   lines( nc[,1], nc[,2], col="red", lwd=2 )
   tet1 <- tet2
   pp < -nc[200,]
```





```
# En este ejemplo tenemos funciones anidadas
KochSnowflakeExample <- function() {  # Función general</pre>
   iterate <- function(T,i){ # Primera función anidada</pre>
  A = T[,1]; B=T[,2]; C = T[,3];
   if (i == 1) {
    d = (A + B)/2; h = (C-d); d = d-(1/3)*h;
    e = (2/3) *B + (1/3) *A; f = (1/3) *B + (2/3) *A;
   if (i == 2) {
    d = B; e = (2/3) *B + (1/3) *C; f = (2/3) *B + (1/3) *A;
   if (i == 3) {
    d = (B + C)/2; h = (A-d); d = d-(1/3)*h;
    e = (2/3) \cdot C + (1/3) \cdot B; f = (1/3) \cdot C + (2/3) \cdot B;
   if (i == 4) {
    d = C; e = (2/3) * C + (1/3) * A; f = (2/3) * C + (1/3) * B;
```

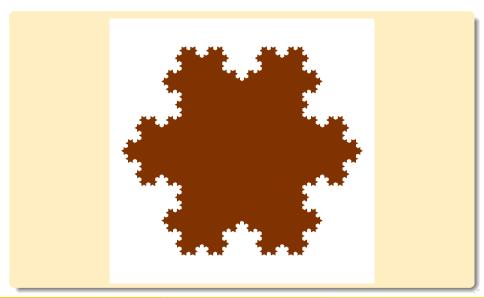
```
if (i == 5) {
    d = (A + C)/2; h = (B-d); d = d-(1/3)*h;
    e = (2/3) *A + (1/3) *C; f = (1/3) *A + (2/3) *C;
  if (i == 6) {
    d = A; e = (2/3) *A + (1/3) *C; f = (2/3) *A + (1/3) *B;
  if (i == 0) {
    d = A; e = B; f = C;
      Tnew = cbind(d,e,f)
    return (Tnew): # Devuelve un triángulo más pequeño.
# Segunda función anidada
  draw <- function(T, col=rgb(0,0,0),border=rgb(0,0,0)) {
      polygon(T[1,],T[2,], col=col, border=border)
```

# \$

```
# Tercera función anidada
Iterate = function(T, v, col=rgb(0,0,0), border=rgb(0,0,0)) {
  for (i in v) T = iterate(T,i):
  draw(T, col=col, border=border);
# Los vértices del triángulo inicial:
A = matrix(c(1,0),2,1);
B = matrix(c(cos(2*pi/3), sin(2*pi/3)), 2,1);
C = \text{matrix}(c(\cos(2*pi/3), -\sin(2*pi/3)), 2, 1);
T0 = cbind(A,B,C):
plot(numeric(0), xlim=c(-1.1,1.1), ylim=c(-1.1,1.1),
   axes=FALSE, frame=FALSE, ann=FALSE);
par(mar=c(0,0,0,0),bg=rgb(1,1,1));
par (usr=c(-1.1,1.1,-1.1,1.1));
```

# €E

```
# Dibujar copo de nieve:
  for (i in 0:6) for (j in 0:6) for (k in 0:6) for (l in 0:6)
Iterate(T0,c(i,j,k,l));
}
# Ejecutamos la función
KochSnowflakeExample()
```



### Índice

- Nociones básicas de R
  - R Studio
  - La consola de R
- Tipos de objetos en R
  - Vectores
  - Matrices y arreglos
  - Listas y factores
  - Data frame
- 3 Lectura/Escritura de datos
  - Leer datos de un archivo
  - Guardar datos
- Como programar en R
- Gráficos con F
- Resumen



## Resumen de comandos y operadores

#### Instrucciones en R:

| Operador        | Símbolo | Operador/<br>Función | Símbolo/<br>Instrucción | Operador/<br>Función | Símbolo<br>Instrucción |
|-----------------|---------|----------------------|-------------------------|----------------------|------------------------|
| Igualdad        | ==      | Suma                 | +                       | Tangente             | tan                    |
| Diferente       | !=      | Resta                | -                       | Máximo               | max                    |
| Menor           | <       | Multiplicación       | *                       | Mínimo               | min                    |
| Menor igual     | <=      | División             | /                       | Rango                | range                  |
| Mayor           | >       | Módulo               | %%                      | Longitud             | length                 |
| Mayor igual     | >=      | División entera      | %/%                     | Sumatoria            | sum                    |
| "y" lógico      | &       | Raíz cuadrada        | sqrt                    | Producto             | prod                   |
| "o" lógico      |         | Log. neperiano       | log                     | Media                | mean                   |
| Negación lógica | !       | Exponencial          | exp                     | Dev. estándar        | sd                     |
|                 |         | Seno                 | sin                     | Varianza             | var                    |
|                 |         | Coseno               | cos                     | Cuantiles            | quantile               |



## Resumen de comandos y operadores

#### Algunas instrucciones:

| Operador           | Símbolo |
|--------------------|---------|
| Núm. filas         | nrow    |
| Núm. columnas      | ncol    |
| Producto matricial | %*%     |
| Transpuesta        | t       |
| Diagonal           | diag    |

## Resumen de comandos y operadores

#### Más instrucciones

| Función         | Operador      | Función         | Operador      |
|-----------------|---------------|-----------------|---------------|
| Como caracter   | as.character  | ¿Es caracter?   | is.character  |
| Como entero     | as.integer    | ¿Es entero?     | is.integer    |
| Como complejo   | as.complex    | ¿Es complejo?   | is.complex    |
| Como numérico   | as.numeric    | ¿Es numérico?   | is.numeric    |
| Como lógico     | as.logic      | ¿Es lógico?     | is.logic      |
| Como factor     | as.factor     | ¿Es factor?     | is.factor     |
| Como data.frame | as.data.frame | ¿Es data.frame? | is.data.frame |
| Como matriz     | as.matrix     | ¿Es matriz?     | is.matrix     |
| Como tabla      | as.table      | ¿Es tabla?      | is.table      |
| Como vector     | as.vector     | ¿Es vector?     | is.vector     |
| Como lista      | as.list       | ¿Es lista?      | is.list       |

