

Actividad 1

Curso de Ciencia de Datos con R

Última vez compilado 13 March, 2023

Abra un archivo de R y responda las siguientes preguntas. En cada respuesta tiene que poner tu respuesta y/o el código de R .

Ejemplo

```
# Tu nombre completo
# Ejercicio 1
# 1. código de R
# 2. código de R
# 3. código de R
# 4. código de R
# 5. código de R
```

Ejercicio 1

Coerción

¿En base a tu conocimiento sobre vectores sin usar ninguna función de R. ¿Cuál es el tipo de vector que se obtiene en los siguientes casos?

1. `c(8, FALSE)`
2. `c("zz", 99)`
3. `c(TRUE, "a")`
4. `c(TRUE, 1L)`
5. `c(8L, 4)`

Ejercicio 2

Vectores

A lo largo de los ejercicios, los vectores están definidos como vectores columnas. En todos los casos realice las operaciones de forma vectorial.

1. Defina dos vectores x , y tal que $x = (3, 2, 2, 4, 5)^T$ e $y = c(1, 6, 5, 5, 9)^T$. Inspeccione el largo de cada vector y el modo usando alguna función de R base.
2. Obtenga $z = 2x_1 + 3x_4 + 5y_2^{\frac{1}{2}}$
3. Obtenga $\sum_{i=1}^5 y_i$, $\sum_{i=1}^5 y_i^2$, $(\sum_{i=1}^5 y_i)^2$, $\sum_{i=1}^5 y_i^{x_i}$.
4. Obtenga la media y mediana tanto de x como de y .
5. Obtenga $\sum_{i=1}^5 x_i y_i \mathbb{1}_{x_i=y_i}$.

Ejercicio 3

Listas

Arme una lista con nombre “lista” con los siguientes componentes:

1. Un vector numérico (numeric) de longitud 4
2. Un caracter (character)
3. Un vector caracter (character - string) de longitud 4
4. Un vector lógico (boolean) de longitud 3
5. El número 1000
6. Seleccione el tercer miembro de la lista
7. Seleccione el tercer elemento del primer miembro de la lista y multiplíquelo por 8
8. ¿Cuál es el valor del segundo elemento del cuarto miembro de la lista?

Ejemplo

```
# Tu nombre completo
# Ejercicio 1
# 1. código de R
# 2. código de R
# 3. código de R
```

Ejercicio 4 (FACTOR)

Sea

```
x <- c("bueno", "muy bueno", "excelente", "bueno")
```

1. Transformar **x** en un factor con los niveles “malo”, “bueno”, “muy bueno”, “excelente” y guardarlo en un objeto llamado **new_x**.
2. Usando la función **ordered** ordenar los niveles del factor **new_x** en orden reverso (excelente, muy bueno, bueno, malo) y guardarlo como un nuevo objeto **new_xx**
3. Ahora usá la función **relevel** para hacer que el primer nivel de **new_x** sea excelente,.
4. Ahora transformá el objeto **x** en un factor con los niveles “excelente”, “muy bueno”, “bueno”, “malo” y que las etiquetas sean “e”, “mb”, “b” y “m”. Nombrar el nuevo objeto como **order_x**

Ejercicio 5 (MATRIX ARRAY)

Generá lo siguiente:

1. Una matriz $M_{5 \times 5}$ cuyos elementos vayan del 1 al 25, donde el primer vector fila de la matriz es (1, 2, 3, 4, 5) y así sucesivamente. Una matriz $J_{5 \times 5}$ donde el primer vector columna es (1, 2, 3, 4, 5), etc.
2. Obtenga $M + J$, MJ , JM
3. Genere $Z = \text{vec}(M)$. Retenga los primeros 10 elementos y calcule $A = ZZ'$ y finalmente obtenga los elementos $A_{i,i}$ con $i = 1, 2, \dots, 10$
4. Dado `x <- c(1:4, 17, 99)` Conviértalo en una matriz de dimensión 3x2 de dos formas: usando la función **matrix** y sin usar la función **matrix**.
5. Genere la siguiente matrix `A <- matrix(c(1, 3, 2, 2, 8, 9), ncol = 3)` y el siguiente vector `a <- c(1, 3, 2)`. Calcule $A \times a$ usando el operador “`%*%`”. Explique la diferencia.
6. Generá un array como el siguiente

```
## , , 1
##
##      [,1] [,2]
## [1,]    1    4
## [2,]    2    5
## [3,]    3    6
##
## , , 2
##
##      [,1] [,2]
## [1,]   10   13
## [2,]   11   14
## [3,]   12   15
##
## , , 3
##
##      [,1] [,2]
## [1,]   20   23
## [2,]   21   24
## [3,]   22   25
```

7. Seleccioná la tercer matriz del array

Ejercicio 6 (DATA FRAME)

1. Creá un data frame con 5 filas y 4 columnas que contengan dos variables categóricas (sexo y color de pelo, inventá los datos) y dos variables numéricas (edad, peso en kg). Nombrá el objeto como `df`
2. ¿Qué tipo de variables hay en cada columna, que función usas para ver eso?
3. Agregá una nueva variable numérica con la altura en cm, usa `cbind()` y guardalo en un nuevo data frame llamado `df2`.
4. Agregá un nuevo registro a tu primer data frame `df` que sea Mujer, de pelo Negro, 28 años y pese 53 kg. Para esto utilizá la función `rbind()`
5. ¿Qué pasó con la estructura de los datos?
6. Usando la función `which.max()` obtenga la observación con la edad máxima usando tu primer data frame `df`.