

Práctica Gestión Datos R - Explicación

Lamberto X. Sánchez Néné

#1

```
vector_edad_descubrimiento <- c(10,25,50,80,97,103,124,153)
mean(vector_edad_descubrimiento)
```

Con el comando “mean” calculamos la media del vector.

#2

```
vector_cantidad_artefactos <- c(1,12,94,3,2,75,203,47)
sum(vector_cantidad_artefactos)
```

Con el comando “sum” calculamos la suma del vector.

#3

```
vector_profundidad_hallazgo <- c(5,21,7,28,10,2,50,31)
m <- which.max(vector_profundidad_hallazgo)
vector_profundidad_hallazgo[m]
```

Con el comando “which.max” destacamos cual es la cifra mayor, dado y reflejado el valor en [m].

#4

```
vector_materiales_encontrados <-
c("Hueso", "Monedas", "Cadáveres", "Ánforas", "Tesoros", "Piedra")
length(vector_materiales_encontrados)
```

Con el comando “length” calculamos cuanto valores diferentes hay en el vector.

#5

```
vector_años_excavaciones <-
c(2001,1989,1981,1973,2014,2022,1963,1956)
length(vector_años_excavaciones)
```

Con el comando “length” calculamos cuanto valores diferentes hay en el vector.

#6

```
matriz1 <- matrix(c(1,2,3,4,5,6,7,8,9,10),  
                  nrow = 5, ncol = 6,  
                  byrow = T)  
rowSums(matriz1)  
which.max(rowSums(matriz1))
```

#7

```
matriz2 <- matrix(c(1,4,6,8,9,12),  
                  nrow = 3, ncol = 4,  
                  byrow = T)  
colMeans(matriz2)  
which.max(colMeans(matriz2))
```

#8

```
matriz3 <- matrix(c(5,9,20,2,17),  
                  nrow = 4, ncol = 5,  
                  byrow = T)  
colSums(matriz3)  
which.max(colSums(matriz3))
```

#9

```
matriz4 <- matrix(c(2,4,6,8,10,14),  
                  nrow = 3, ncol = 4,  
                  byrow = T)  
colMeans(matriz4)  
which.min(colMeans(matriz4))
```

#10

```
matriz5 <- matrix(c(2,5,7,10,12,15),  
                  nrow = 3, ncol = 4,  
                  byrow = T)  
mode <- function(matriz5) {  
  return(as.numeric(names(which.max(table(matriz5)))))  
}  
mode(matriz5)
```

Creamos una matriz, establecemos las columnas y filas, y generamos los códigos que necesitamos. RowSums (Suma de filas), ColMeans (Media de columnas), ColSums (Suma de columnas). Con Which.max y Which.min destacamos respectivamente el valor máximo y mínimo de lo que pidamos.

#11

```
set.seed(123)
n_registros <- 10

encontrados <- sample(c("Carmona", "Santiponce", "Valencina"),
n_registros, replace = TRUE)
tipo_artefaco <- sample(c("monedas", "ánforas", "cadáveres"),
n_registros, replace = TRUE)
fecha_descubrimiento <- sample(1900:2000, n_registros, replace
= TRUE)
descripción <- sample(c("Buen estado", "Estado neutro", "Mal
estado"), n_registros, replace = TRUE)

registro_artefactos <- data.frame(
  encontrados = encontrados,
  tipo_artefaco = tipo_artefaco,
  fecha_descubrimiento = fecha_descubrimiento,
  descripción = descripción
)
print(registro_artefactos)
```

#12

```
set.seed(123)
n_registros <- 5

realizadas <- sample(c("Grupo 1", "Grupo 2", "Grupo 3"),
n_registros, replace = TRUE)
sitio <- sample(c("Huelva", "Sevilla", "Cádiz"), n_registros,
replace = TRUE)
fecha_inicio <- sample(2000:2010, n_registros, replace = TRUE)
fecha_final <- sample(2011:2020, n_registros, replace = TRUE)

excavaciones_equipo <- data.frame(
  realizadas = realizadas,
  sitio = sitio,
  fecha_inicio = fecha_inicio,
  fecha_final = fecha_final
)
print(excavaciones_equipo)
```

#13

```
set.seed(123)
n_registros <- 10

descubiertos <- sample(c("Carmona", "Valencina", "Santiponce"),
n_registros, replace = TRUE)
sitio <- sample(c("Tumba", "Horno", "Tierra"), n_registros,
replace = TRUE)
edad <- sample(c(20, 40, 75, 90, 100), n_registros, replace =
TRUE)
sexo <- sample(c("Hombre", "Mujer"), n_registros, replace =
TRUE)
huesos <- sample(c("100% encontrados", "75% encontrados", "50%
encontrados", "25% encontrados", "0% encontrados"), n_registros,
replace = TRUE)

datos_esqueletos <- data.frame(
  descubiertos = descubiertos,
  sitio = sitio,
  edad = edad,
  sexo = sexo,
  huesos = huesos
)
print(datos_esqueletos)
```

#14

```
set.seed(123)
n_registros <- 7
ubicación <- sample(c("Yacimiento 1", "Yacimiento
2", "Yacimiento 3", "Yacimiento 4"), n_registros, replace =
TRUE)
sitio <- sample(c("Castillo", "Teatro", "Hogar"), n_registros,
replace = TRUE)
latitud <- sample(c(4:20), n_registros, replace = TRUE)
longitud <- sample(c(30:100), n_registros, replace = TRUE)
altitud <- sample(c(300:500), n_registros, replace = TRUE)

ubicaciones_geográficas <- data.frame(
  ubicación = ubicación,
  sitio = sitio,
  latitud = latitud,
  longitud = longitud,
  altitud = altitud
)
print(ubicaciones_geográficas)
```

Creamos una semilla con “set.seed”, establecemos el número de registros con “n_registros”, ahora establecemos los campos que queramos mencionar “x <- sample (c(nombre1, nombre2, ...), n_registros, replace = TRUE”. Se pueden incluir tantos valores textuales entre comillas y números. Más adelante otorgamos un nombre al “data.frame”, y establecemos relación entre todos los campos, “ubicación = ubicación, etc”. Finalmente con el comando “print” podemos ver reflejado todo el proceso.