

PROBABILIDADES Y ESTADÍSTICA (C)
PRÁCTICA 0
INTRODUCCIÓN A R

1. a) Sea

```
x <- c(1,2,3)
y <- c(6,5,4).
```

Predecir el resultado de correr cada una de las siguientes líneas y después chequear las respuestas.

```
x * 2
x * y
x[1] * y[2]
1/x
(1:10) * x[2]
rep(c(1,1,2), times = 2)
```

- b) ¿Qué ocurre si se multiplican 2 vectores de distinta longitud?

2. a) Crear un vector llamado tratamiento con coordenadas A , B y C de manera que A aparezca 20 veces, B 18 y C 22.

- b) Definir un vector J que sea una secuencia de 5 a 100 con un incremento de 5 y luego sumar la primera y la octava coordenadas.

3. Sea $q = (3, 0, 1, 6)$ y $r = (1, 0, 2, 4)$

- a) Calcular el producto escalar entre q y r
- b) Construir una matriz cuyas filas sean q y r . LLamarla a .
- c) Construir una matriz cuyas filas sean q y r . LLamarla b .
- d) Calcular el producto de a y b .
- e) Calcular la inversa de $a * b$.
- f) Reemplazar la primera columna de a por una columna de unos.
- g) Extraer la segunda columna de b .

4. Calcular la suma de los cuadrados de los números naturales del 1 al 100 usando R.

5. Considerar el conjunto de datos `mtcars` en R. (Ver `help(mtcars)`)

- a) ¿Qué autos tienen 4 velocidades?
- b) ¿Qué autos hay en el subconjunto `mtcars[mtcars$disp > 150 & mtcars$mpg > 20,]`?
- c) ¿Qué autos tienen 4 velocidades y transmisión manual?
- d) Hallar la cantidad media de millas por galón de los autos con 2 carburadores.

6. El conjunto de datos `arbolado-en-espacios-verdes.csv` contiene datos de 2011 de todos los árboles de los espacios verdes de la ciudad de Buenos Aires. Los datos pueden encontrarse aquí, junto con una descripción de los mismos .

- a) ¿Cuántas filas tiene el conjunto de datos? ¿Cuántas columnas? ¿Cuáles son los nombres de las columnas?
 - b) Calcular la altura promedio de los árboles de la ciudad en 2011.
 - c) ¿Cuántos árboles había en 2011 en la plaza Arenales?
 - d) Construya un `data.frame` llamado `arboles_favoritos` que contenga sólo las filas correspondientes a su espacio verde favorito. Puede utilizar el comando `sort` para ordenar alfabéticamente un vector de caracteres.
 - e) Utilizando el comando `unique` averigüe los nombres de los árboles presentes en su espacio verde favorito en 2011.
7. a) Definir una función llamada `elegir_alumnos` que elija al azar m alumnos de un aula de n para exponer ejercicios en el pizarrón, asumiendo que los alumnos están numerados de 1 a n .
- b) Simular el experimento de elegir 3 alumnos al azar de un aula de 80 alumnos para exponer ejercicios en el pizarrón.
- c) Repetir 10 veces el experimento del ítem 7b y guardar los resultados en una matriz llamada `alumnos_elegidos`
8. Algunos dicen que las últimas horas a bordo del Titanic estuvieron marcadas por la guerra de clases, otros sostienen que estuvieron caracterizadas por la caballerosidad de los varones. Los datos TITANIC3 del paquete PASWR2 contienen información sobre los pasajeros del Titanic, incluyendo clase, sexo y si sobrevivieron o no, entre otras características.
- (a) Determine la proporción de sobrevivientes por clase.
- (b) Calcule la proporción de sobrevivientes por clase y sexo. ¿Quién tuvo una tasa más alta de supervivencia: los varones de 1ra clase o las mujeres de 3ra?
- (c) Hacer un histograma con los datos correspondientes a la edad de los pasajeros con 20 intervalos. ¿Qué observa? Si dividimos el rango de las edades en 20 grupos de 5 años cada uno: 0-5, 6-10, ..., 75-80. ¿Cuál era el rango de edades más numeroso en el Titanic? Y el menos numeroso? ¿Cómo trata R los valores faltantes al hacer el histograma?
- (d) ¿Cuál era la edad de la mujer más joven que sobrevivió?
- (g) En su opinión y basándose en estos datos ¿fue la guerra de clases, la caballerosidad de los varones o una combinación de ambas lo que caracterizó las últimas horas del Titanic?
9. El conjunto de datos CARS2004 del paquete PASWR2 contiene datos de automóviles en Europa del año 2004. (ver `help(CARS2004)`)
- a) Calcular cantidad total de autos en cada país.
 - b) Calcular la tasa de mortalidad de automovilistas para cada país como el número total de muertes de automovilistas dividido el número total de autos.
 - c) Hacer un gráfico de barras que indique la tasa de mortalidad en accidentes de tránsito de cada país de la Unión Europea. Ordene las barras de forma creciente.

- d)* ¿Qué país tiene la menor tasa de mortalidad de automovilistas y qué país la más baja?
- e)* Haga un gráfico de cantidad total de autos vs población. ¿Cómo describiría la relación?
- f)* Haga un gráfico de cantidad total de autos vs tasa de mortalidad de automovilistas. ¿Cómo describiría la relación?