

Ejercicios Tema 1. Segunda Entrega.

Pau Vives, Harold Cruz, Samuel de Paúl

PARTE I

1. Calcula la tasa de asesinatos por cada 100000 habitantes para cada estado y almacénala en un objeto llamado `murder_rate`. Luego, usa operadores lógicos para crear un vector lógico llamado `low` que nos dice qué entradas de `murder_rate` son inferiores a 1.

```
#Importamos los datos
data(murders)

#Calculamos la tasa
murder_rate <- murders$total / murders$population * 100000

#entradas inferiores a 1:
low <- murder_rate < 1
```

2. Ahora usa los resultados del ejercicio anterior y la función `which` para determinar los índices de `murder_rate` asociados con valores inferiores a 1.

```
#Directamente buscamos en qué posiciones están los TRUE del vector lógico del
#ejercicio anterior
ind = which(low == TRUE)
ind
```

```
## [1] 12 13 16 20 24 30 35 38 42 45 46 51
```

3. Usa los resultados del ejercicio anterior para indicar los nombres de los estados con tasas de asesinatos inferiores a 1.

```
#Escogemos de la columna de los estados aquellos dados por la lista de índices del
#ejercicio anterior
murders$state[ind]
```

```
## [1] "Hawaii"      "Idaho"        "Iowa"         "Maine"
## [5] "Minnesota"   "New Hampshire" "North Dakota" "Oregon"
## [9] "South Dakota" "Utah"         "Vermont"      "Wyoming"
```

4. Ahora extiende el código de los ejercicios 2 y 3 para indicar los estados del noreste con tasas de asesinatos inferiores a 1. Sugerencia: use el vector lógico predefinido `low` y el operador lógico `&`.

```
murders$state[low & murders$region == "Northeast"]
```

```
## [1] "Maine"      "New Hampshire" "Vermont"
```

5. En un ejercicio anterior, calculamos la tasa de asesinatos para cada estado y el promedio de estos números. ¿Cuántos estados están por debajo del promedio?

```
#Calculamos la tasa media
murder_mean <- mean(murder_rate)
murder_mean
```

```
## [1] 2.779125
```

```
#Creamos un vector lógico que nos dice, de las entradas de murder_rate,
#cuáles están por debajo de la tasa media, y utilizamos sum para ver de
#cuantos estados se trata
below_mean <- murder_rate < murder_mean
sum(below_mean)
```

```
## [1] 27
```

6. Usa la función `match` para identificar los estados con abreviaturas AK, MI e IA. Sugerencia: comienza definiendo un índice de las entradas de `murders$abb` que coinciden con las tres abreviaturas. Entonces usa el operador `[` para extraer los estados.

```
ind <- match(c("AK", "MI", "IA"), murders$abb)
murders$state[ind]
```

```
## [1] "Alaska" "Michigan" "Iowa"
```

7. Utiliza el operador `%in%` para crear un vector lógico que responda a la pregunta: ¿cuáles de las siguientes son abreviaturas reales: MA, ME, MI, MO, MU?

```
#Creamos el vector lógico como se indica tomamos los índices correspondientes
abbs <- c("MA", "ME", "MI", "MO", "MU")
ind = abbs %in% murders$abb
abbs[ind]
```

```
## [1] "MA" "ME" "MI" "MO"
```

```
#De hecho podemos reutilizar la expresión del ejercicio anterior para saber
#a qué estados corresponden
murders$state[match(abbs[ind], murders$abb)]
```

```
## [1] "Massachusetts" "Maine" "Michigan" "Missouri"
```

8. Extiende el código que usaste en el ejercicio 7 para averiguar la única entrada que no es una abreviatura real. Sugerencia: usa el operador `!`, que convierte `FALSE` a `TRUE` y viceversa, y entonces `which` para obtener un índice.

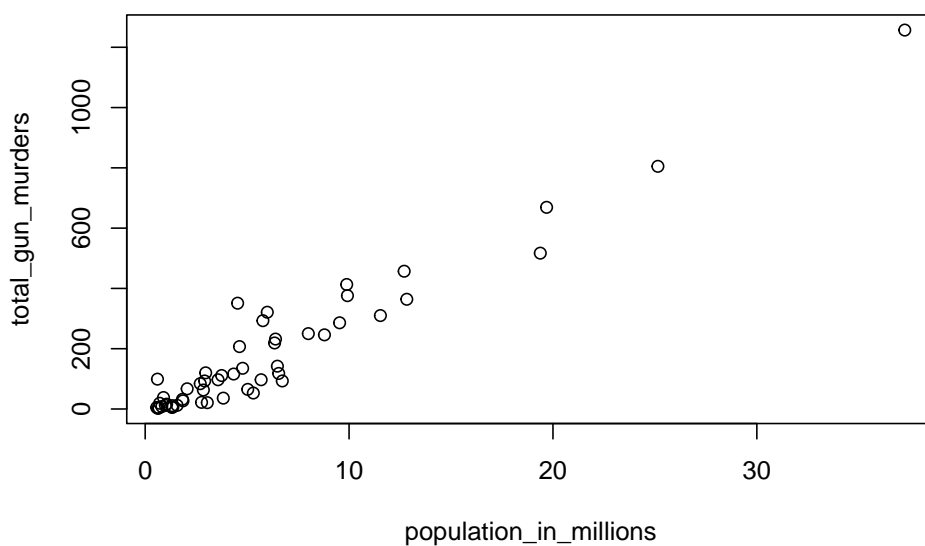
```
abbs <- c("MA", "ME", "MI", "MO", "MU")
ind = abbs %in% murders$abb
which(!ind)
```

```
## [1] 5
```

PARTE II:

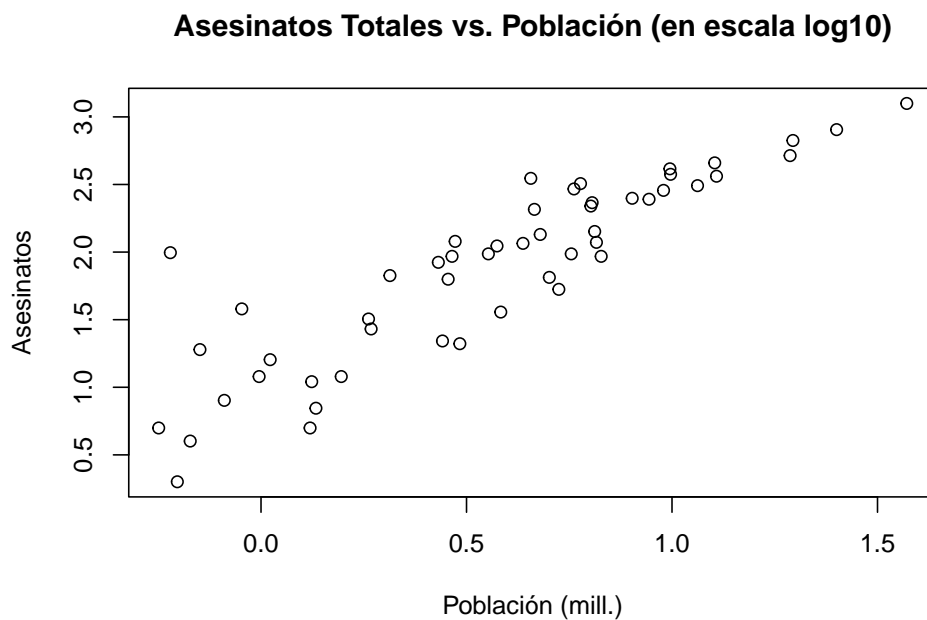
1. Hicimos un gráfico de asesinatos totales versus población y notamos una fuerte relación. No es sorprendente que los estados con poblaciones más grandes hayan tenido más asesinatos.:

```
population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total
plot(population_in_millions, total_gun_murders)
```



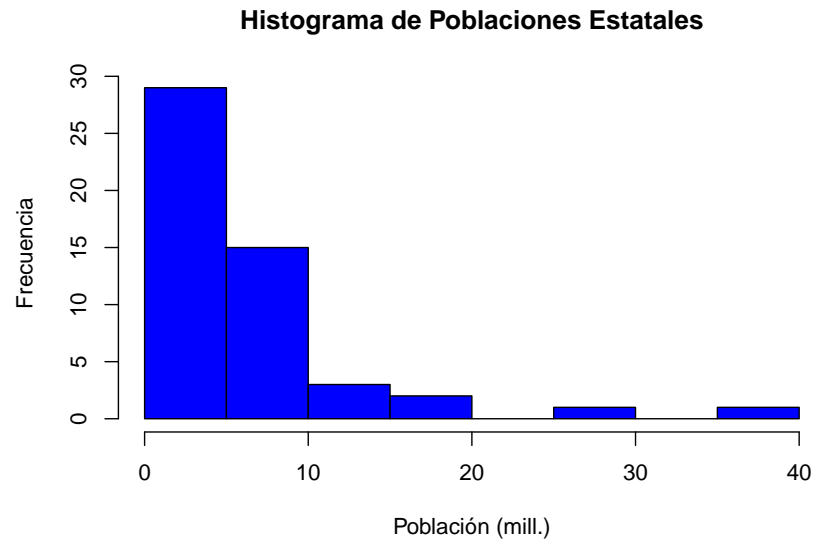
Recuerda que muchos estados tienen poblaciones inferiores a 5 millones y están agrupados. Podemos obtener más información al hacer este gráfico en la escala logarítmica. Transforma las variables usando la transformación \log_{10} y luego crea un gráfico de los resultados.

```
plot(log10(population_in_millions), log10(total_gun_murders),
     main = "Asesinatos Totales vs. Población (en escala log10)",
     xlab = "Población (mill.)", ylab = "Asesinatos")
```



2. Crea un histograma de las poblaciones estatales.

```
hist(population_in_millions, col = "blue",  
main = "Histograma de Poblaciones Estatales", xlab = "Población (mill.)", ylab = "Frecuencia")
```



3. Genere diagramas de caja de las poblaciones estatales por región.

```
#Añadimos la columna de población en millones al dataset  
murders$population_mill <- population_in_millions  
  
#Representamos el boxplot  
boxplot(population_mill~region, data = murders, col = cm.colors(4),  
main = "Boxplot de Población por Regiones", xlab = "Regiones", ylab = "Población (mill.)")
```

