

# Parcial1

Jose Alejandro Zapata Ortiz - Jhon Freddy Castro Madrid

2022-03-24

## R Markdown

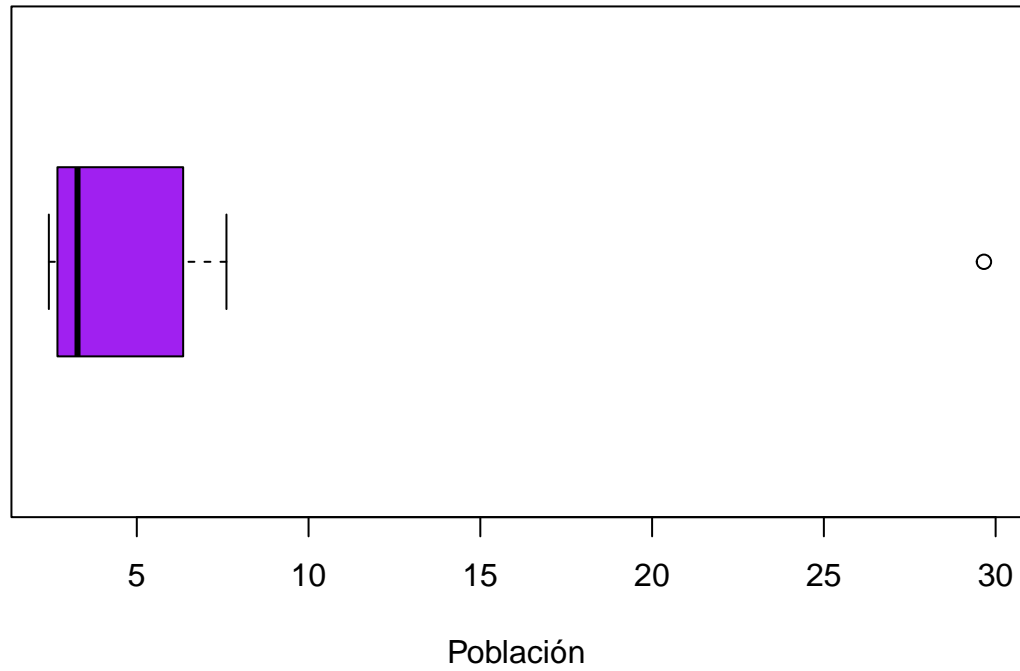
1. La tabla mostrada contiene valores de población, en cientos de miles, de las diez ciudades más pobladas de cuatro países en el año 1967:

```
Argentina <- c(29.66,7.61,6.35,4.10,3.80,2.75,2.70,2.69,2.51,2.44)
EEUU <- c(77.81,35.50,24.79,20.02,16.70,9.39,9.38,8.76,7.63,7.50)
Holanda <-c(8.68,7.31,6.02,2.64,1.75,1.72,1.51,1.42,1.31,1.29)
Japon <- c(110.21,32.14,18.88,16.38,13.37,11.92,10.71,7.80,7.70,7.00)
```

a.Construya un box-plot e identifique los puntos extremos. ¿Cuáles son las características más sobresalientes?  
¿Hay outliers?

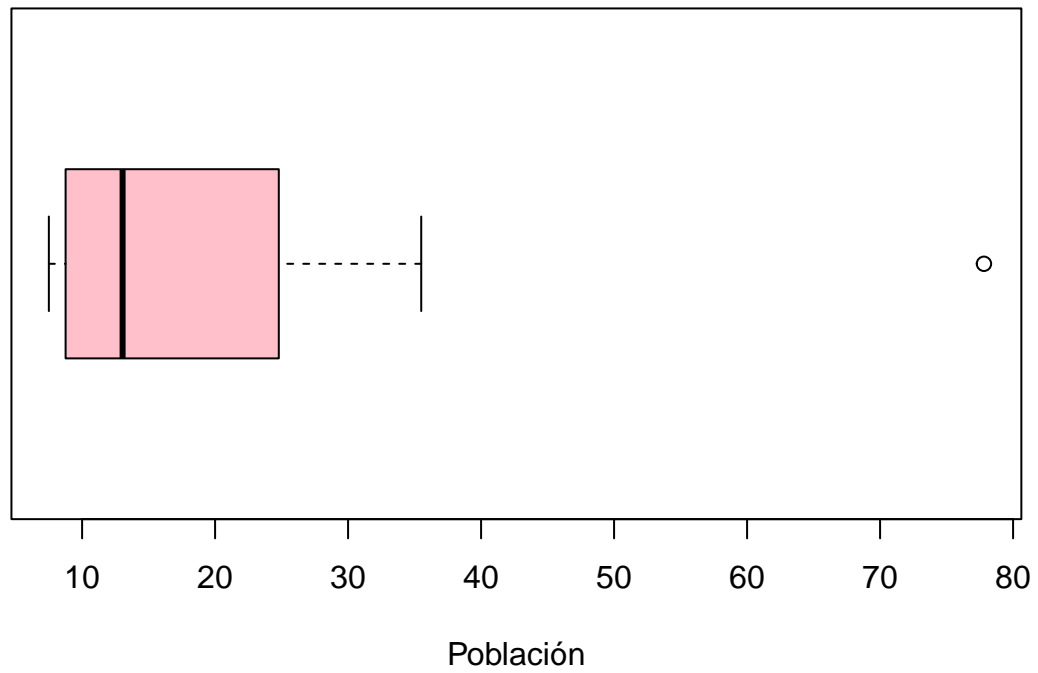
```
boxplot(Argentina, col="purple", main = "Población de ciudades de Argentina",
        xlab = "Población",Xlim=c(0,30), horizontal = TRUE)
```

## Población de ciudades de Argentina



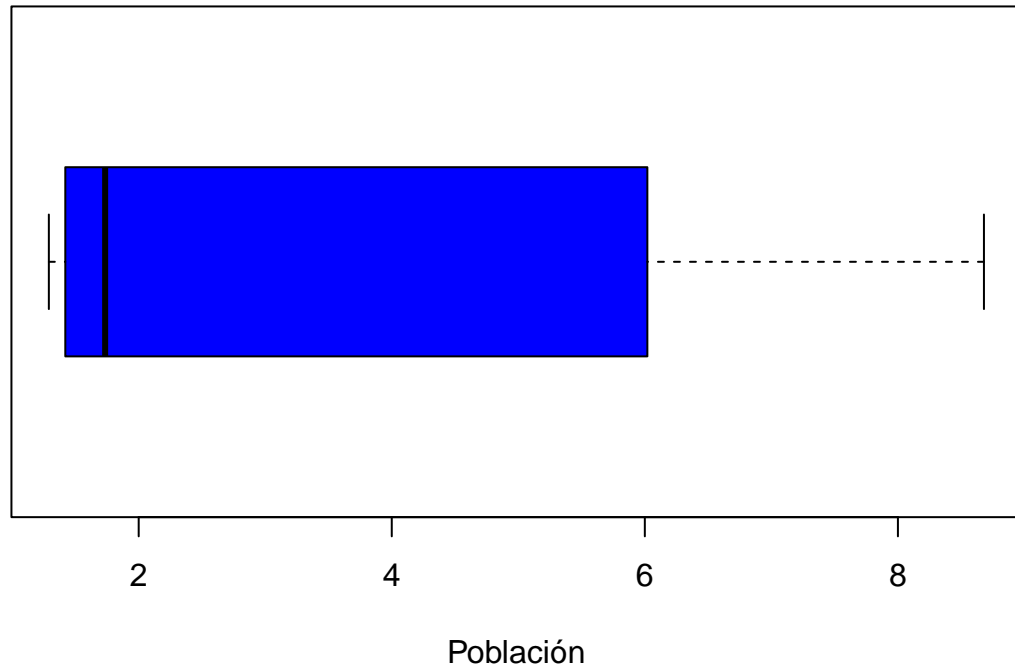
```
boxplot(EUU, col="pink", main = "Población de ciudades de EEUU",  
        , xlab = "Población", horizontal = TRUE)
```

## Población de ciudades de EEUU



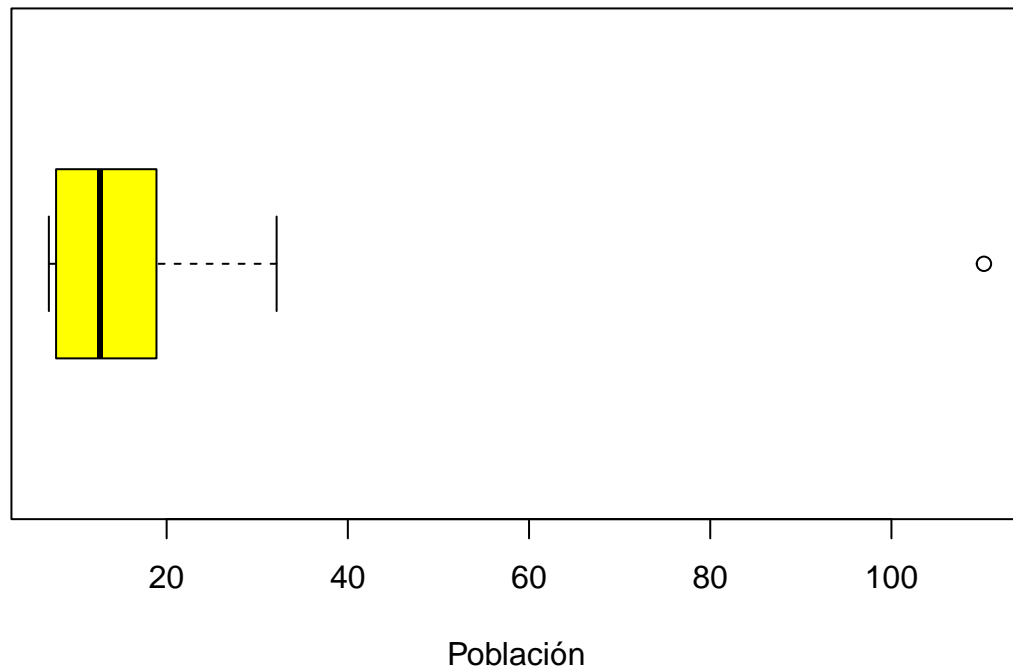
```
boxplot(Holanda, col="blue", main = "Población de ciudades de Holanda",  
        xlab = "Población", horizontal = TRUE)
```

## Población de ciudades de Holanda



```
boxplot(Japon, col="yellow", main = "Población de ciudades de Japon",  
        xlab = "Población", horizontal = TRUE)
```

## Población de ciudades de Japon



R/ a. Se puede identificar que 3 de las 4 graficas cuentan con una simetria positiva, ya que los datos se concentran en la parte inferior de la distribución, donde la poblacion de las ciudades de Japon es la unica que no muestra ese comportamiento ya que tiene un comportamiento mas normal que las demas, ya que en este se observa una simetria normal, ademas se identifica que las poblaciones de las ciudades de Japon, EEUU y Argentina cuentan con datos atipicos, para ser mas especifico uno por cada pais de los .

b. Compare los centros de cada población, sus dispersiones y su simetría. ¿Cuál es el país más homogéneamente habitado? ¿Cuáles son las características más sobresalientes? ¿Hay outliers?

R/ b. Teniendo en cuenta los diagramas de caja y bigotes se puede concluir que:

La población de muestra de Japon son los unicos que parecen tener una simetria normal, los datos son centrados y solo tiene un dato atipico el cual es de 110,21

Los datos de Argentina tienen una simetria positiva, su centro esta en el primer cuantil, tiene un solo dato atipico que es de 29.66

La grafica de los datos de Holanda tiene una simetria positiva, es la que tiene la caja más grande este nos indica que tiene datos muy alejados de su mediana y no posee datos atipicos.

Los datos de la población de Estados Unidos tiene una simetria positiva, tiene su centro mas a la derecha de la mediana, tiene un dato atipico de 77.81.

2. Avianca se encuentra estudiando la situación de venta de puesto en el vuelo Armenia -- Bogotá puesto que ha identificado que el 2% de las personas que reservaron puesto, no se presentan a la hora del embarque. Teniendo en cuenta que el avión A320 utilizado en esta ruta tiene una capacidad para 180 pasajeros, la empresa decide vender 181 tiquetes. ¿Cuál es la probabilidad de que todas las personas que llegan a embarque puedan tener su silla asegurada y no se presenten quejas por "sobre-vender" sillas ante la aeronáutica civil? Asuma que las llegadas de los pasajeros siguen una distribución Poisson.

```
lambda <- 180 * 0.98
q <- 181
resp_pois <- ppois(180,lambda)
print(paste("El porcentaje de que las personas que llegan embarque puedan tener su silla asegurada: ",
```

```
## [1] "El porcentaje de que las personas que llegan embarque puedan tener su silla asegurada: 62.552"
```

3.En la planta envasadora de Coca-Cola FEMSA ubicada en Tocancipá se realiza el embotellado de la Coca-Cola presentación de 400 mL (en promedio), con una desviación estándar de 5 mL (mililitros). Si se sabe por experiencia que este proceso de embotellado sigue una distribución normal, y también se sabe que todo producto con más de 415 mL es declarado como NO CONFORME. Determine el porcentaje de las botellas de gaseosa que son declaradas producto NO CONFORME.

```
media <- 400
ds <- 5
res <- pnorm(415,media,ds,lower.tail = FALSE)

print(paste("El porcentaje de que las botellas que son declaradas
NO CONFORME es de: ",round(res*100,4),'%'))
```

```
## [1] "El porcentaje de que las botellas que son declaradas\n NO CONFORME es de: 0.135 %"
```