

Guía 1

Probabilidad y Estadística

Ingeniería Industrial - Facultad de Ciencias Agrarias

Tema 1. Estadística - Uso de Rstudio

Estadística. Concepto

Estadística es la ciencia que crea, desarrolla y aplica métodos para la toma de decisiones frente a la incertidumbre, partiendo de información empírica y calculando los riesgos.

Campo de Acción

En cualquier actividad humana, la incertidumbre está presente, por lo que la Estadística brinda metodología aplicable a todas las áreas del conocimiento humano. Con el desarrollo de la Ciencia Estadística, que es relativamente reciente, se han ido perfeccionando técnicas y métodos aplicables a diferentes ramas del conocimiento humano. En ese sentido existen métodos generales y otros se han desarrollado especialmente para su aplicación en las Ciencias Sociales, Médicas, Económicas, Biológicas entre otras, y fundamentalmente ha habido un gran desarrollo de diseños y métodos estadísticos aplicables a la Experimentación Agropecuaria e Ingeniería.

Aplicaciones

En el campo de las Ciencias Biológicas y especialmente en las Agronómicas e Industriales, la incertidumbre se incrementa por el hecho de que los factores que afectan los fenómenos de interés son innumerables y muchas veces inmanejables. En esas condiciones la Estadística brinda metodología para la obtención de información empírica confiable y para el análisis adecuado de dicha información a fin de tomar decisiones con sustento científico.

Ramas de la Estadística

La Estadística puede dividirse en tres grandes ramas:

- Estadística Descriptiva: brinda metodología adecuada para obtener información y describir el comportamiento de un conjunto de individuos.
- Estadística Inferencial: permite, diseñar la obtención de información en un subconjunto representativo del total de individuos bajo estudio, y a partir de dicha información extraer conclusiones aplicables a la totalidad de los individuos.
- Estadística Experimental: incluye metodología para la obtención y análisis de información de experimentos, que involucran situaciones en las que el investigador interviene y modifica las condiciones naturales del fenómeno en estudio.

Ejercicio 1

Ingresar los datos correspondientes al consumo de bebida gaseosa a nivel mundial.

```
gaseosas <- read.csv("~/data/gaseosas.csv", sep=";")
gaseosas
```

```
##           pais consumo
## 1 Arabia Saudita   79.5
## 2 Uruguay         87.1
## 3 Alemania        76.3
## 4 Argentina       131.1
## 5 Bélgica         88.5
## 6 Perú            55.9
## 7 Irlanda         74.4
## 8 Colombia        65.8
## 9 Noruega         91.2
## 10 Chile          121.3
## 11 Estados Unidos 112.1
## 12 Brasil         89.9
## 13 México         119.4
```

Indicar cuáles son los países de mayor y menor consumo de gaseosa.

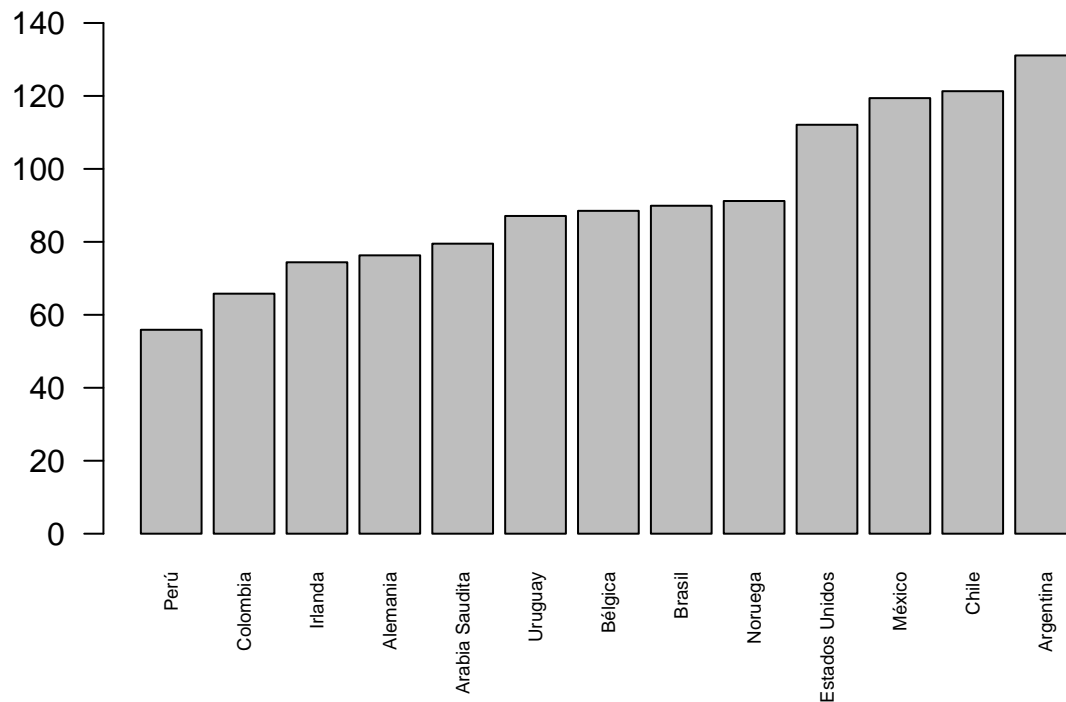
Ordenar los datos de manera ascendente.

```
data <- gaseosas[order(gaseosas$consumo),]
data
```

```
##           pais consumo
## 6 Perú        55.9
## 8 Colombia    65.8
## 7 Irlanda     74.4
## 3 Alemania    76.3
## 1 Arabia Saudita 79.5
## 2 Uruguay     87.1
## 5 Bélgica     88.5
## 12 Brasil     89.9
## 9 Noruega     91.2
## 11 Estados Unidos 112.1
## 13 México     119.4
## 10 Chile      121.3
## 4 Argentina   131.1
```

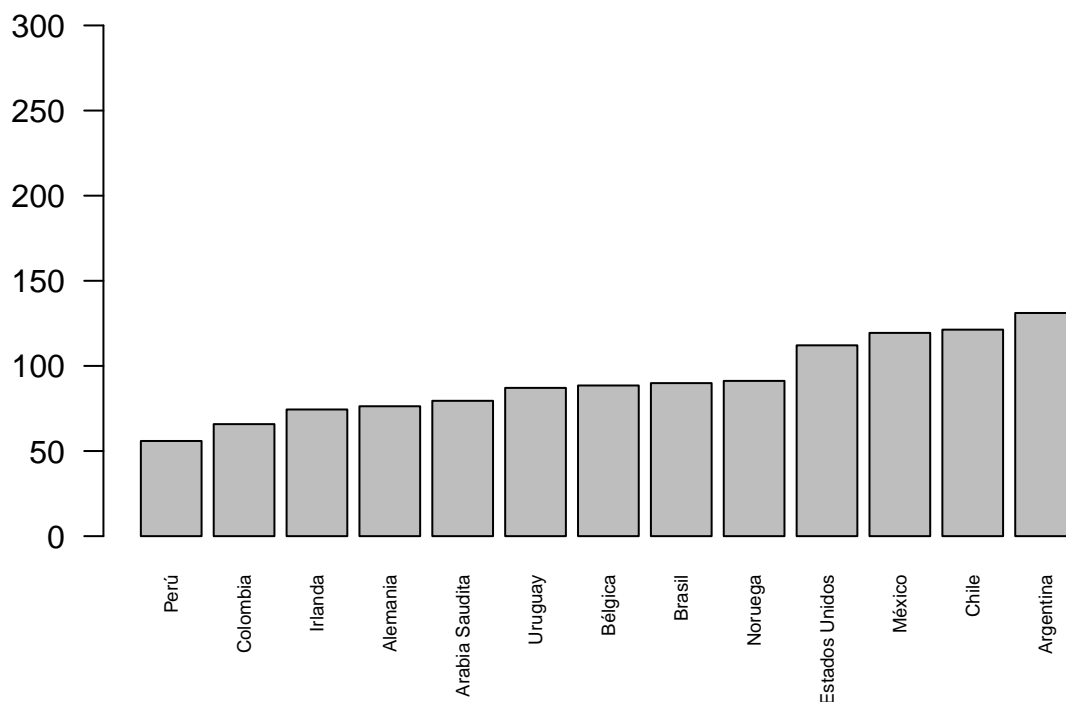
Elaborar un gráfico de barras por países ¿Qué conclusiones puede realizar con respecto al consumo de gaseosa en los diferentes países? ¿Es posible visualizar grupos de países de similar consumo? ¿Cuáles serían esos grupos?

```
barplot(data$consumo,
        ylim = c(0, 140),
        names.arg = data$pais,
        las = 2,
        cex.names = 0.6)
```



modificar la escala del eje de 0 a 300 ¿Puede observar lo mismo que en punto anterior?

```
barplot(data$consumo,
        ylim = c(0, 300),
        names.arg = data$pais,
        las = 2,
        cex.names = 0.6)
```



Transformar la variable Consumo aplicando raíz cuadrada y logaritmo (log 10).

Logaritmo base 10

```
logaritmo <- log10(data$consumo)
data2 <- cbind(data, logaritmo)
data2
```

```
##      pais consumo logaritmo
## 6      Perú    55.9  1.747412
## 8    Colombia    65.8  1.818226
## 7     Irlanda    74.4  1.871573
## 3     Alemania    76.3  1.882525
## 1 Arabia Saudita    79.5  1.900367
## 2      Uruguay    87.1  1.940018
## 5      Bélgica    88.5  1.946943
## 12     Brasil    89.9  1.953760
## 9      Noruega    91.2  1.959995
## 11 Estados Unidos  112.1  2.049606
## 13      México    119.4  2.077004
## 10      Chile    121.3  2.083861
## 4      Argentina   131.1  2.117603
```

Raíz cuadrada

```

raiz_cuadrada <- sqrt(data$consumo)
data3 <- cbind(data2, raiz_cuadrada)
data3

```

```

##           pais consumo logaritmo raiz_cuadrada
## 6          Perú    55.9  1.747412    7.476630
## 8      Colombia    65.8  1.818226    8.111720
## 7        Irlanda    74.4  1.871573    8.625543
## 3        Alemania    76.3  1.882525    8.734987
## 1  Arabia Saudita    79.5  1.900367    8.916277
## 2          Uruguay    87.1  1.940018    9.332738
## 5          Bélgica    88.5  1.946943    9.407444
## 12         Brasil    89.9  1.953760    9.481561
## 9          Noruega    91.2  1.959995    9.549869
## 11 Estados Unidos  112.1  2.049606   10.587729
## 13          México  119.4  2.077004   10.927031
## 10          Chile  121.3  2.083861   11.013628
## 4          Argentina 131.1  2.117603   11.449891

```

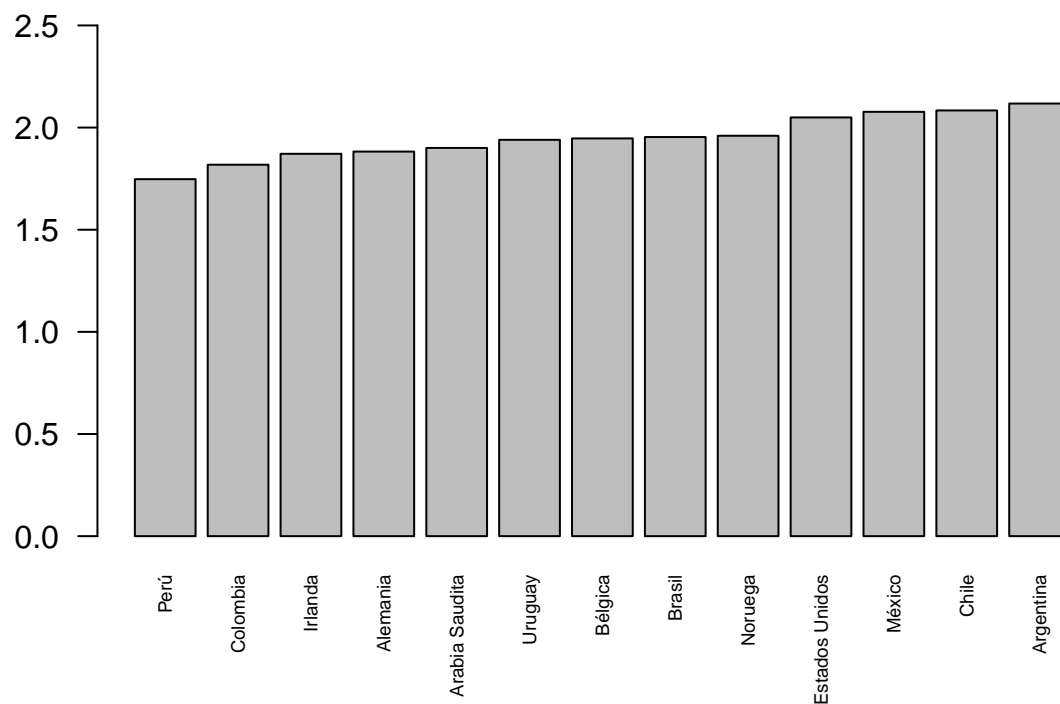
Realizar un solo gráfico de barras (por países) con la variable consumo y sus transformaciones

¿Qué se visualiza al aplicar una transformación? ### Logaritmo 10

```

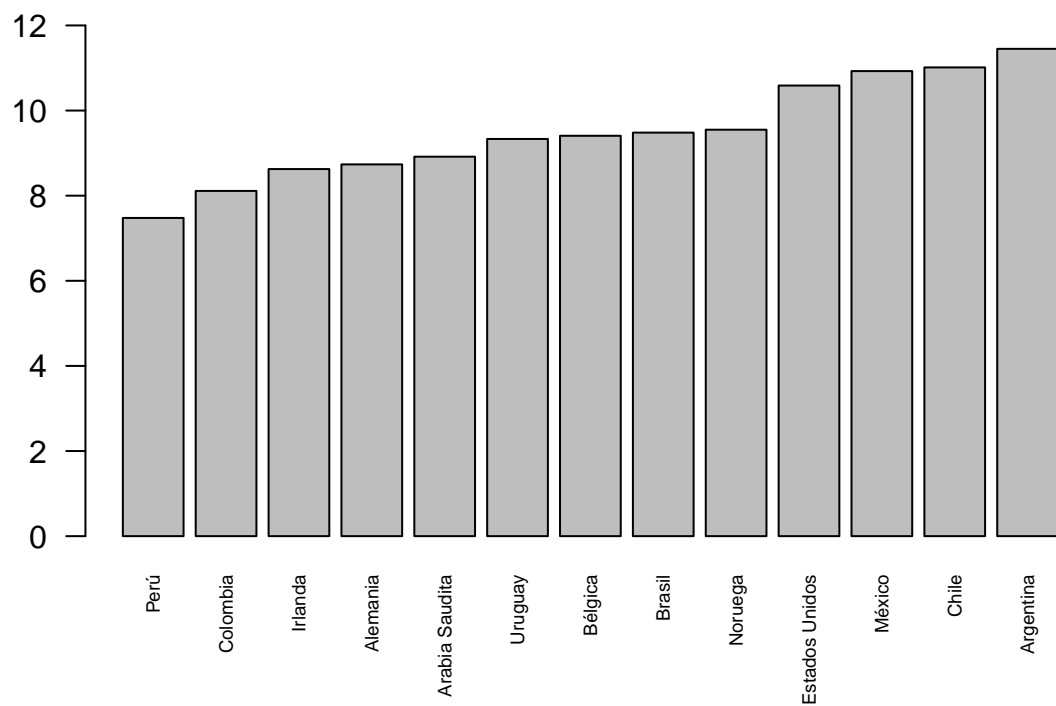
barplot(logaritmo,
        ylim = c(0, 2.50),
        names.arg = data$pais,
        las = 2,
        cex.names = 0.6)

```



Raiz cuadrada

```
barplot(raiz_cuadrada,  
        ylim = c(0, 12),  
        names.arg = data$pais,  
        las = 2,  
        cex.names = 0.6)
```



Interpretar los gráficos de cada variable por separado (consumo, log10, raíz) y concluir.