



International Center for Tropical Agriculture
Since 1967 Science to cultivate change

Análisis de datos agricultura

3 Julio 2018

Juan Camilo Rivera

j.c.rivera@cgiar.org

Hugo Dorado

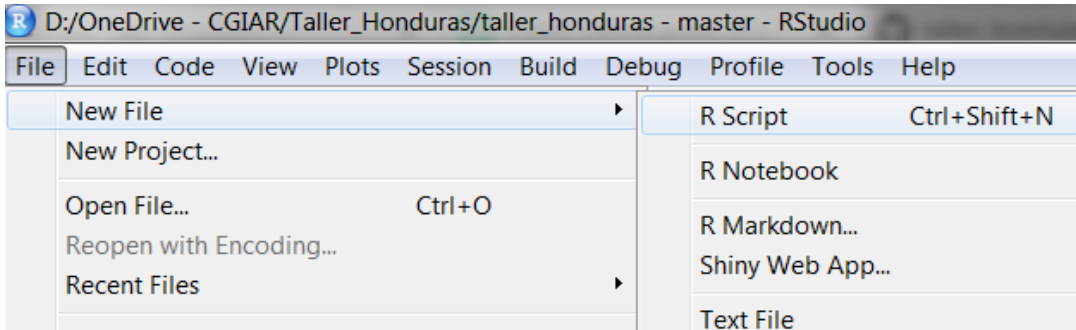
h.a.dorado@cgiar.org



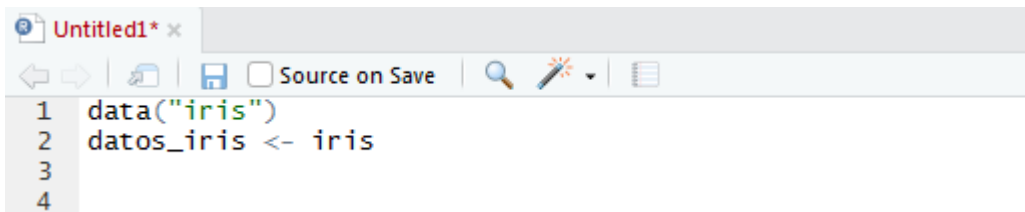
CIAT is a CGIAR Research Center

Antes de empezar

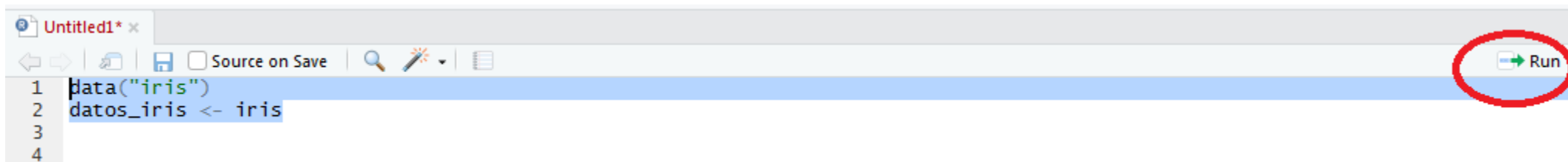
1. Abrir R Studio File >- New File >- R Script



2. Digitar en script.



3. Seleccionar y dar click en run.



Variables

CUALITATIVAS

Nominales:

Si sus valores no se pueden ordenar, ejemplos:

- Sexo (Masculino, Femenino)
- Color ojos (Verde, Azul, Gris, Negro),
- Fumar (Si/No)

Escala Nominal.

Ordinales:

Si sus valores se pueden ordenar, ejemplos:

- Grado de satisfacción (Excelente, Bueno, Regular, Malo)
- Intensidad de dolor (Intenso, leve)

Escala ordinal

CUANTITATIVAS

Discretas:

Si toma valores enteros.

Número de hijos, Número de carros.

Escala de la razón:

El valor cero refleja ausencia de la característica.

Continuas:

Si entre dos valores, son posibles infinitos valores intermedios.

Altura, Temperatura, Duración de una batería, Peso(kg).

Escala de intervalo:

El valor 0 es un valor arbitrario, no implica la no presencia de una característica.

Código en R

```
> names(datos_iris)
[1] "Sepal.Length" "Sepal.width"  "Petal.Length" "Petal.width"  "Species"
> summary(datos_iris)
```

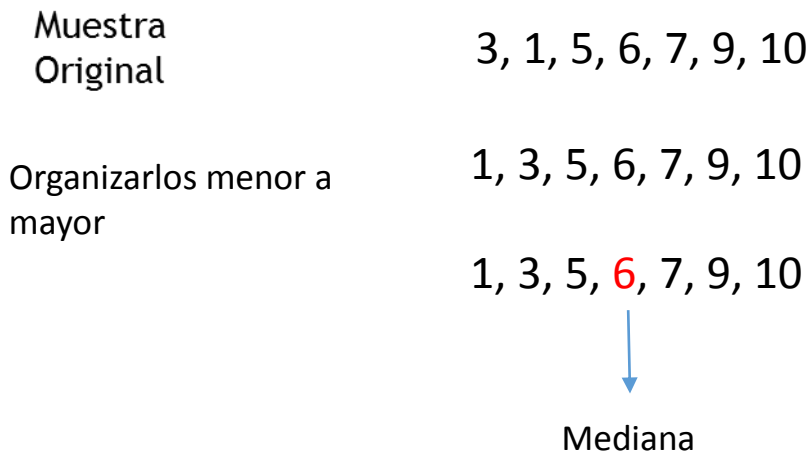
Sepal.Length	Sepal.width	Petal.Length	Petal.width	Species
Min. :4.300	Min. :2.000	Min. :1.000	Min. :0.100	setosa :50
1st Qu.:5.100	1st Qu.:2.800	1st Qu.:1.600	1st Qu.:0.300	versicolor:50
Median :5.800	Median :3.000	Median :4.350	Median :1.300	virginica :50
Mean :5.843	Mean :3.057	Mean :3.758	Mean :1.199	
3rd Qu.:6.400	3rd Qu.:3.300	3rd Qu.:5.100	3rd Qu.:1.800	
Max. :7.900	Max. :4.400	Max. :6.900	Max. :2.500	

Medidas descriptivas

- Media
- Mediana
- Moda
- Varianza
- Desviación estandar

Mediana

Es el valor que separa la mitad de las observaciones ordenadas de menor a mayor.



¿Que pasa si el tamaño de la muestra es **impar**?

$$\begin{array}{c} 20, 2, 6, 12 \\ 2, 6, 12, 2 \\ \downarrow \quad \downarrow \\ (6+12) \div 2 \\ \mathbf{9} \end{array}$$

Media

Promedio aritmético de las observaciones.

$$\bar{x} = \sum_{i=0}^n \frac{x_i}{n}$$

$$\begin{array}{c} 5, 15, 3, 2, 25 \\ \underbrace{\hspace{1.5cm}} \\ n = 5 \\ \bar{x} = \frac{5+15+3+2+25}{5} = 10 \end{array}$$

Moda

Es el valor de la variable que más veces se repite, es decir, aquella cuya frecuencia absoluta es mayor.

1,2,2,3,3,3,3,1,12,4
 $\underbrace{\hspace{1.5cm}}$
3

Unimodal

1,2, 2, 2, 3,3,3,12,13
 $\underbrace{\hspace{1.5cm}} \quad \underbrace{\hspace{1.5cm}}$
2 3

Bimodal o
Multimodal

Varianza

- Es el promedio del cuadrado de las distancias entre cada observación y la media aritmética del conjunto de observaciones. Que tan dispersos están los datos

$$V = \frac{\sum_{i=1}^{i=n} (X_i - \bar{X})^2}{n-1}$$

Desviación Estandar

Que tan dispersos están los datos respecto a la media.

$$S = \sqrt{V}$$

5, 15, 3, 2, 25

$$V = \frac{(5-10)^2 + (15-10)^2 + (3-10)^2 + (2-10)^2 + (25-10)^2}{5}$$

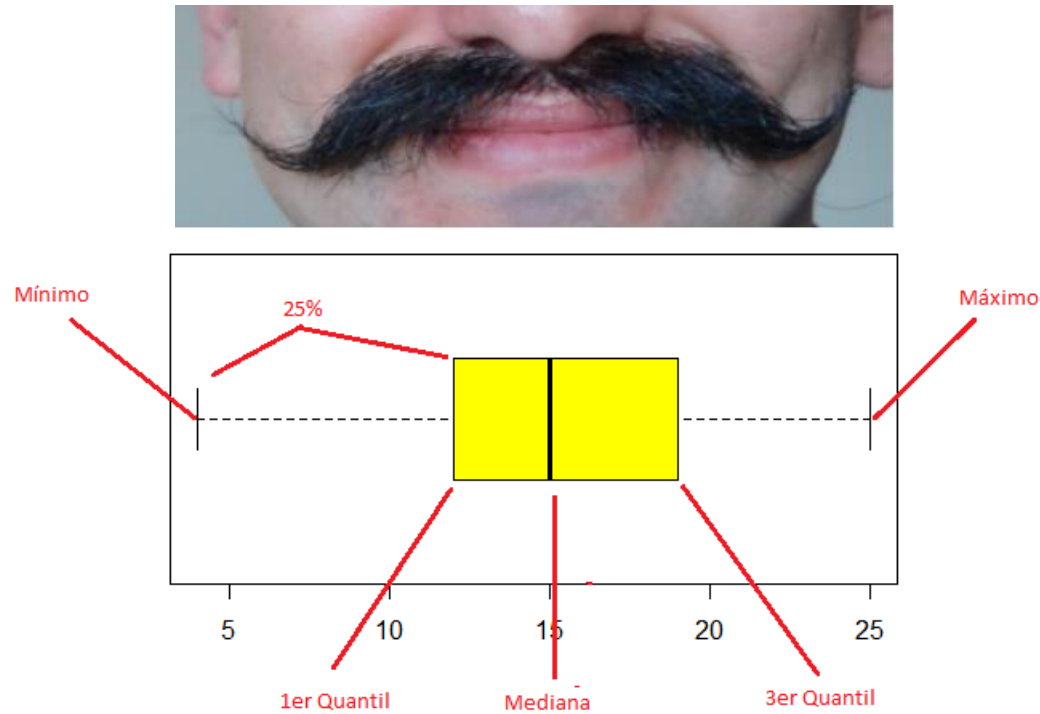
$$V = 97$$

$$S = \sqrt{97}$$

$$S = 9.8488$$

Cuantiles

Son puntos que dividen la muestra en cuatro partes iguales.



- Ejemplo:

Datos: 6, 7, 15, 36, 39, 40, 41, 42, 43, 47, 49

2 Quantil (Mediana): 6, 7, 15, 36, 39, 40, 41, 42, 43, 47, 49

1 Quantil: 6, 7, 15, 36, 39, 40, 41, 42, 43, 47, 49

3 Quantil: 6, 7, 15, 36, 39, 40, 41, 42, 43, 47, 49

Total Quantiles: 6, 7, 15, 36, 39, 40, 41, 42, 43, 47, 49

Codigo en R

```
#Mediana
x <- c(3,1,5,6,7,9,10)
mediana <- median(x)
mediana

#Mediana muestra par
y <- c(20,2,6,12)
mediana <- median(y)

#Promedio
promedio_x <- mean(x)
promedio_y <- mean(y)

#Moda
z <- c(1,2,2,2,3,3,3,1,12,4)

moda <- function(x) {
  t <- table(x)
  return(as.numeric(names(t)[t == max(t)]))
}

moda_z <- moda(z)
moda_z

#Cuantiles
cuantiles <- quantile(datos_iris$Sepal.Length)
cuantiles
```

Thank you!



WE'RE PROUD TO
HAVE CELEBRATED 50 YEARS
OF AGRICULTURAL RESEARCH
FOR DEVELOPMENT

International Center for Tropical Agriculture - CIAT

Headquarters and Regional Office
for South America and the Caribbean

+57 2 445 0000

Km 17 Recta Cali-Palmira
A.A. 6713, Cali, Colombia

✉ ciat.cgiar.org

🌐 ciat.cgiar.org



CIAT is a CGIAR Research Center