



Centro Internacional de Agricultura Tropical
Desde 1967 *Ciencia para cultivar el cambio*

Gráficos descriptivos en R

Julio 2018

Autores

Hugo Andrés Dorado

Juan Camilo Rivera

h.a.dorado@cgiar.org , j.c.rivera@cgiar.org



El CIAT es un Centro de Investigación de CGIAR

Gráficos estadísticos con R

Gráficos base

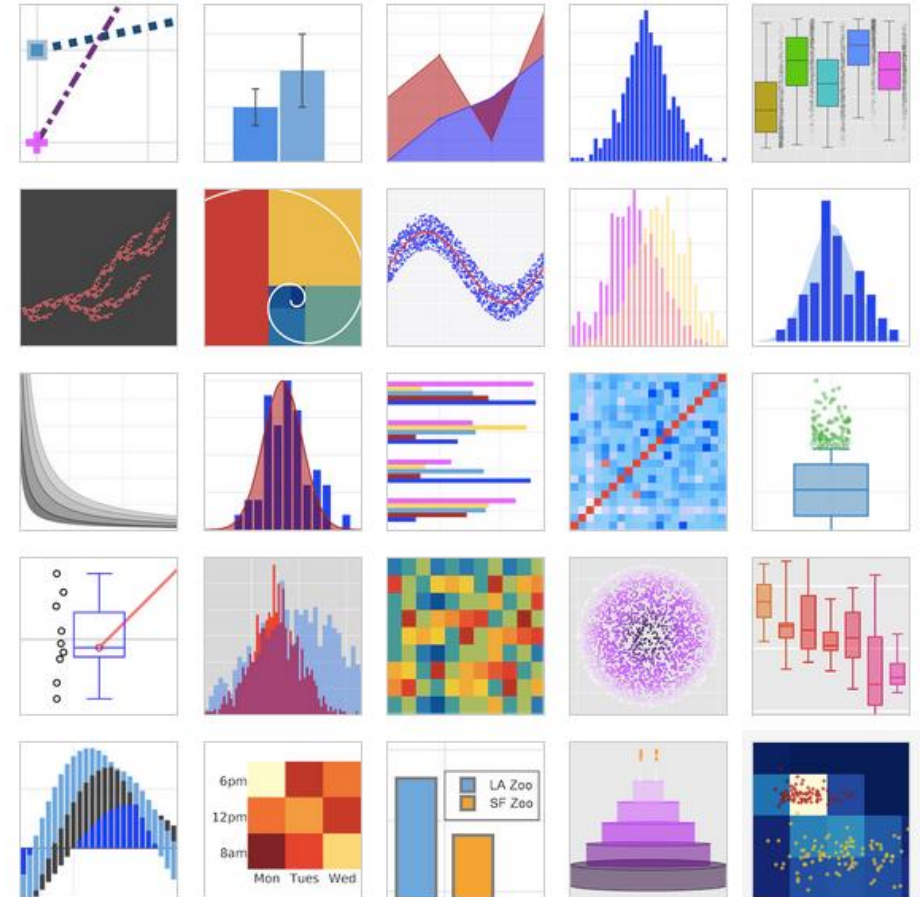
- Solo producen un único resultado por gráfico.
- No requiere paquetes adicionales.

ggplot

- Combinación de funciones que proporcionan los componentes del gráfico.
- Permiten crear varios datos simultáneos divididos por una o varias categorías.
- Ideal para presentaciones, informes, exploración.

lattice

- Utilizan una estructura matricial de paneles definida a partir de una forma
- Permiten crear varios datos simultáneos divididos por una o varias categorías.
- Ideal para publicaciones.



Img. [datasciencecentral.com](https://www.datasciencecentral.com)

Características de gráficos

William Playfair el pionero de la estadística gráfica, realizó su investigación basado en los siguientes principios.

1. El método gráfico es una forma de simplificar lo tedioso y lo complejo.
2. Las personas ocupadas necesitan alguna clase de ayuda visual.
3. Un gráfico es más accesible que una tabla
4. El método gráfico es concordante con los ojos.
5. El método gráfico ayuda al cerebro, ya que permite entender y memorizar mejor.

Definición de un gráfico 'bueno' según Winer 1990

“Un gráfico fuertemente bueno muestra todo lo que queremos conocer sólo con mirarlo.”

“Un gráfico débilmente bueno nos muestra lo que necesitamos conocer observándolo, una vez sepamos como mirarlo.”

Fuente: <https://cran.r-project.org/doc/contrib/grafi3.pdf>

Principios y elementos de un gráfico

Principios

- Entendibilidad
- Claridad
- Consistencia
- Eficiencia
- Necesidad
- Confiabilidad

Elementos de un gráfico.

- Título Principal
- Descripción del Grafico
- Región de Datos y Símbolos
- Eje Horizontal y Escala
- Eje Vertical y Escala
- Descriptores de Señales y marcas

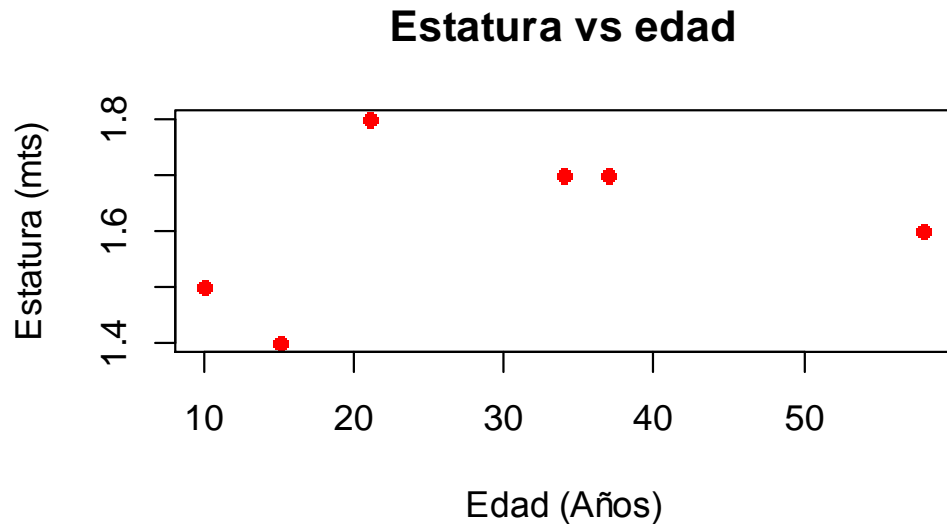
Gráficos en R paquete base

Gráfico de dispersión

```
edad      <- c(21,34,10,15,37,58)
```

```
estatura  <- c(1.8,1.7,1.5,1.4,1.7,1.6)
```

```
plot(edad,estatura,col="red",pch=16,main= 'Estatura vs edad', xlab =  
'Edad (Años)', ylab = 'Estatura (mts)')
```



Símbolos para plot (pch)

0	1	2	3	4	
□	○	△	+	×	
5	6	7	8	9	
◇	▽	⊠	✱	⬢	
10	11	12	13	14	
⊕	⊗	⊞	⊠	⊡	
15	16	17	18	19	
■	●	▲	◆	●	
20	21	22	23	24	25
●	●	■	◆	▲	▼

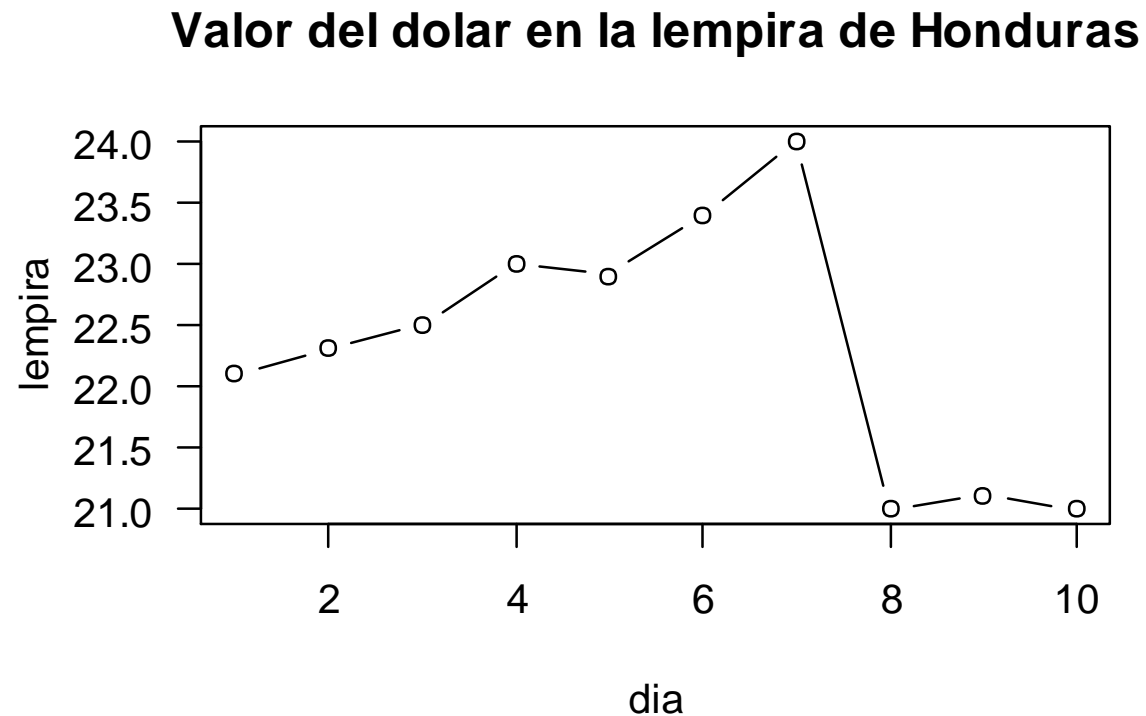
Fuente: <http://www.sthda.com/english/wiki/r-plot-pch-symbols-the-different-point-shapes-available-in-r>

Gráficos de series temporales

```
dia <- c(1:10)
```

```
lempira <- c(22.1,22.3,22.5,23,22.9,23.4,24,21,21.1,21)
```

```
plot(dia,lempira,type='b',las=1,main='Valor del dolar en la lempira de Honduras')
```

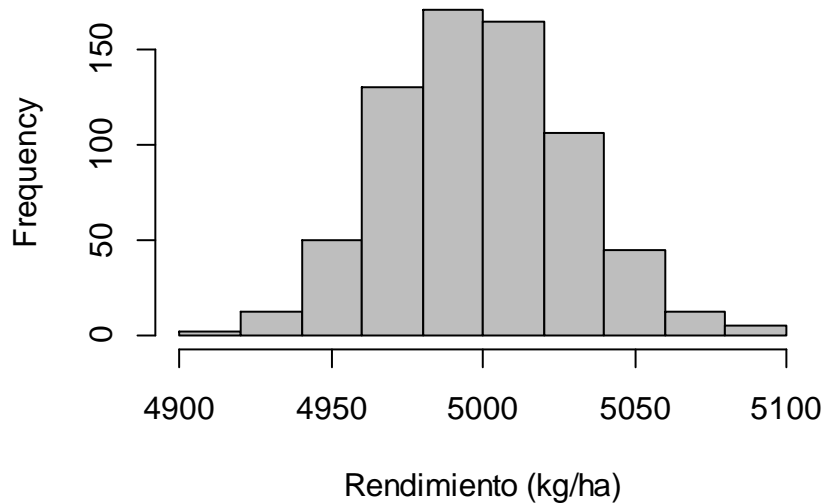


Histograma de frecuencia y diagrama de barras

```
x <- rnorm(700,mean = 5000, sd = 30 )  
hist(x,xlab = 'Rendimiento (kg/ha)',main = 'Histograma  
rendimiento' ,col='gray')
```

Variables cuantitativas

Histograma rendimiento en arroz



```
barplot(tabcultivos, ylim=c(0,25), main = 'Diagrama de barras de  
cultivos',ylab='Frecuencia')
```

```
box()
```

Variables cualitativas

Diagrama de barras de cultivos

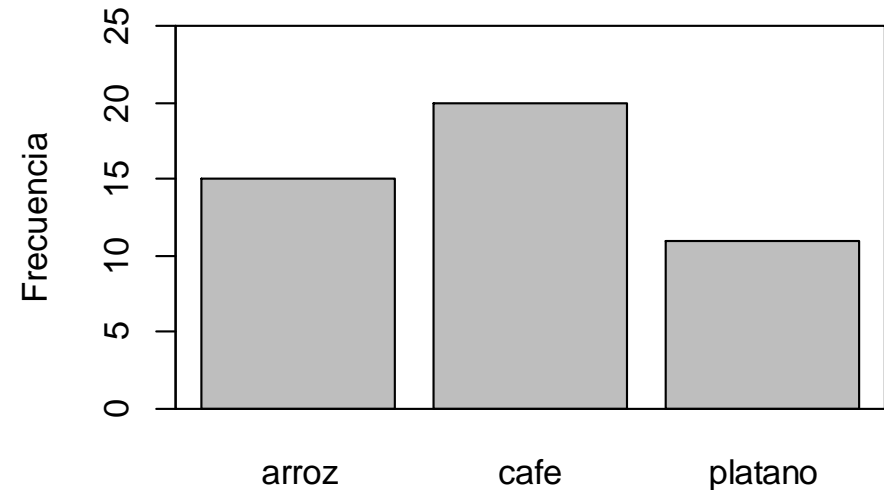
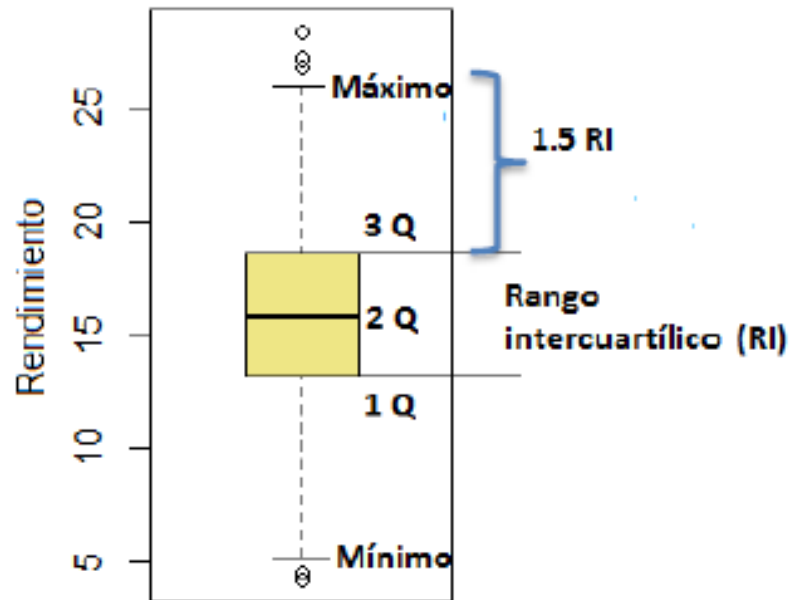
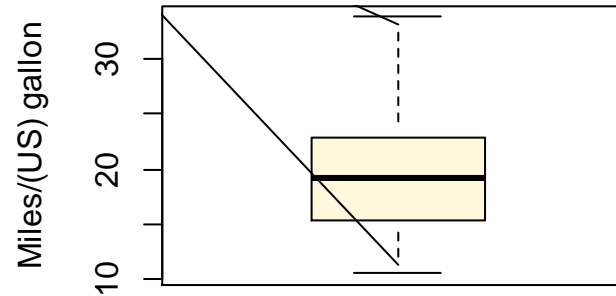


Gráfico de boxplot



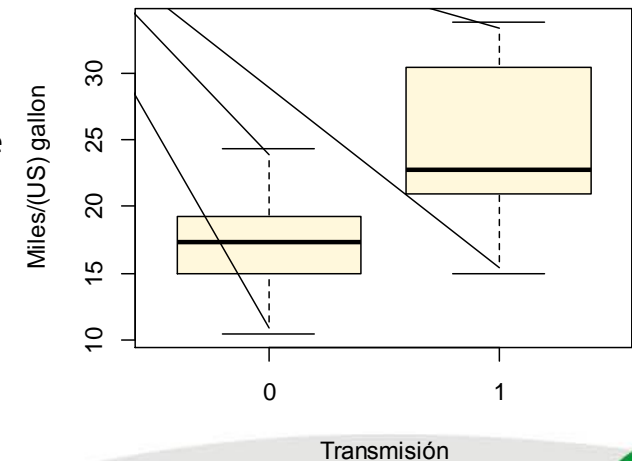
Consumo de combustible



```
boxplot(mtcars$mpg,main='Consumo de combustible',  
        ylab='Miles/(US) gallon',col='cornsilk')
```

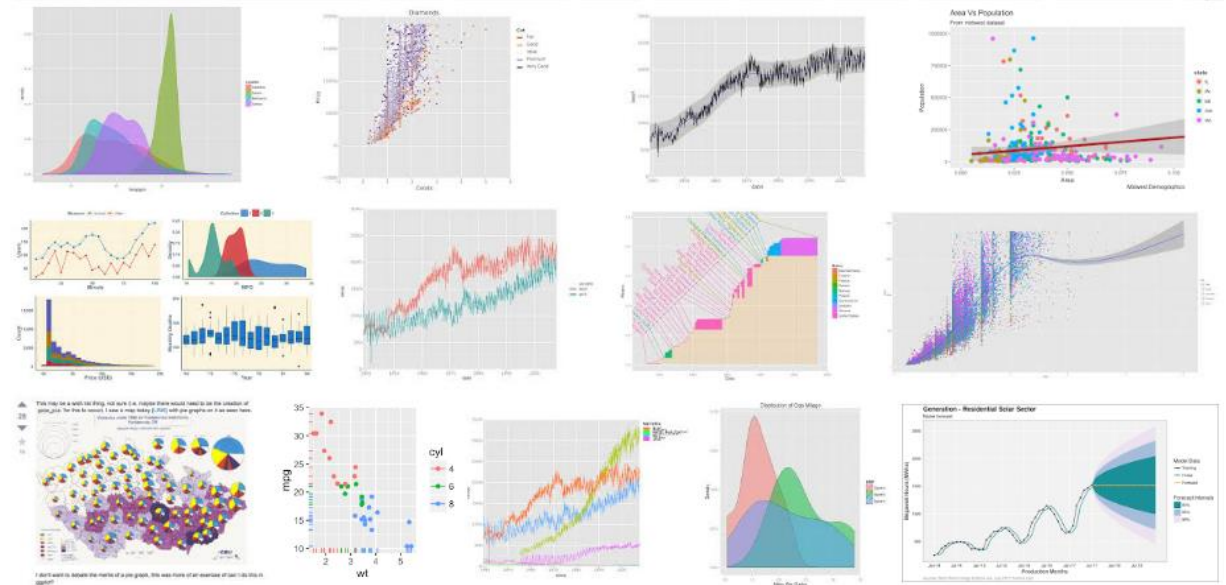
```
boxplot(mpg~am,data=mtcars,main='Consumo de  
combustible por tipo de transmisión',  
        ylab='Miles/(US)  
gallon',col='cornsilk',xlab='Transmisión')
```

Consumo de combustible



Gráficos con ggplot

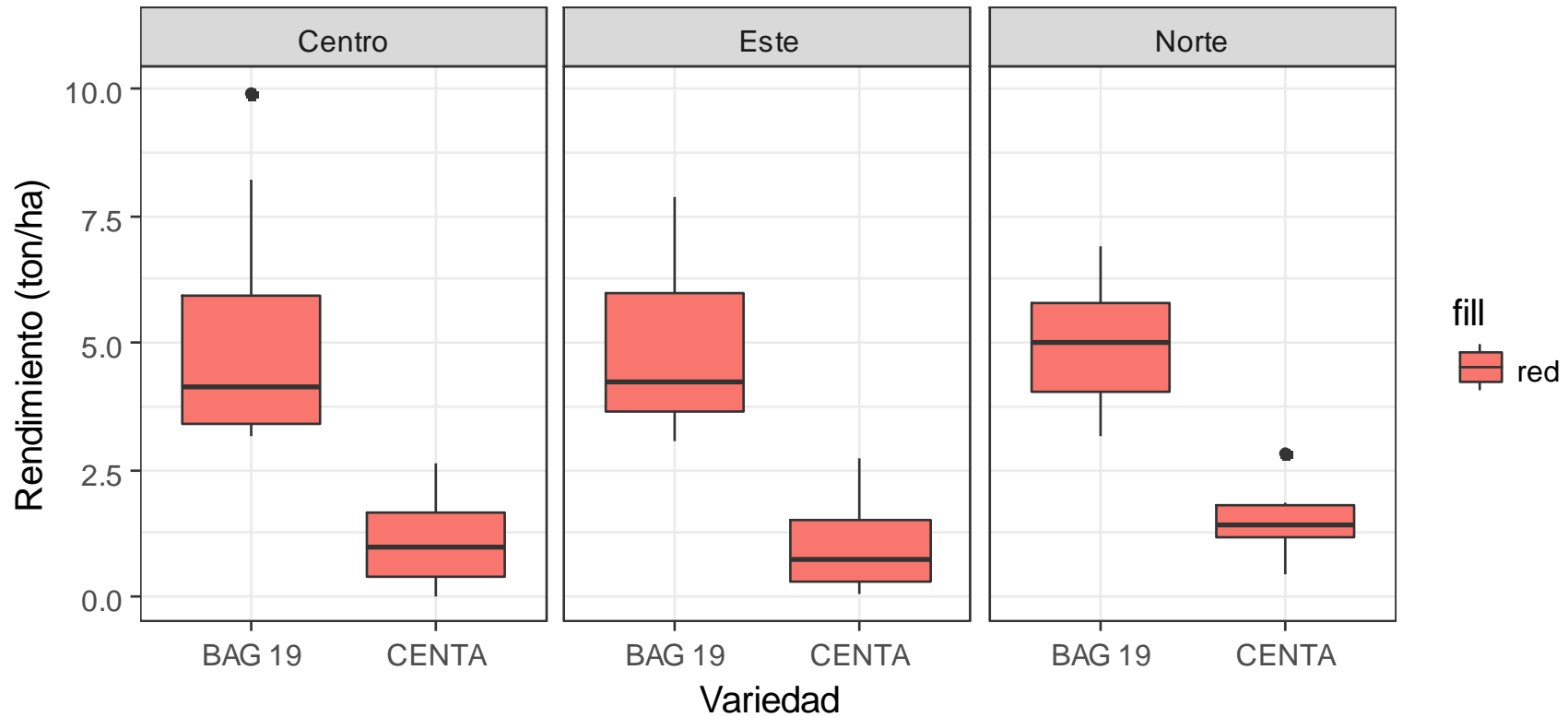
- El orden de los comando es inspirado en una estructura gramatical de gráficos (Wilkinson, 2005)
- Especificación de los gráficos con un alto nivel de abstracción.
- Muy flexible
- Distintos temas y apariencias.
- Muchos usuarios activos



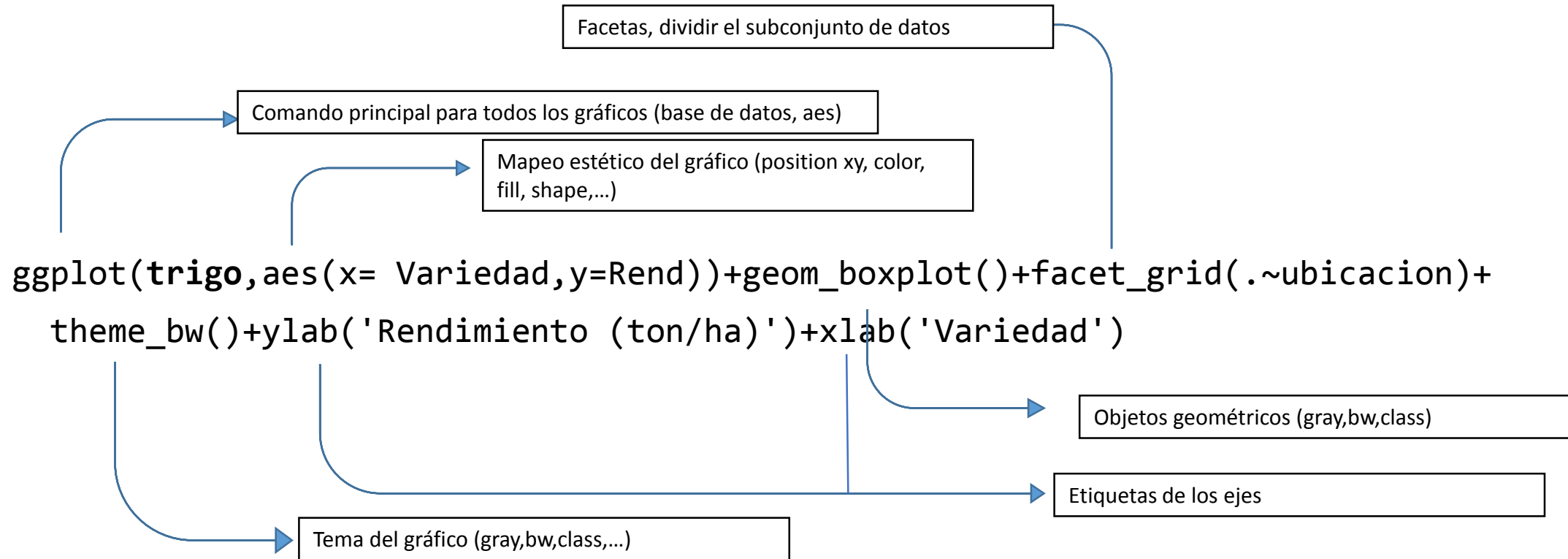
Fuente: <https://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html>

Gráficos con ggplot

```
ggplot(trigo,aes(x= Variedad,y=Rend,fill='red'))+geom_boxplot()+facet_grid(.~ubicacion)+  
  theme_bw()+ylab('Rendimiento (ton/ha)')+xlab('Variedad')
```



Gramática de graficos



<https://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html#scales>

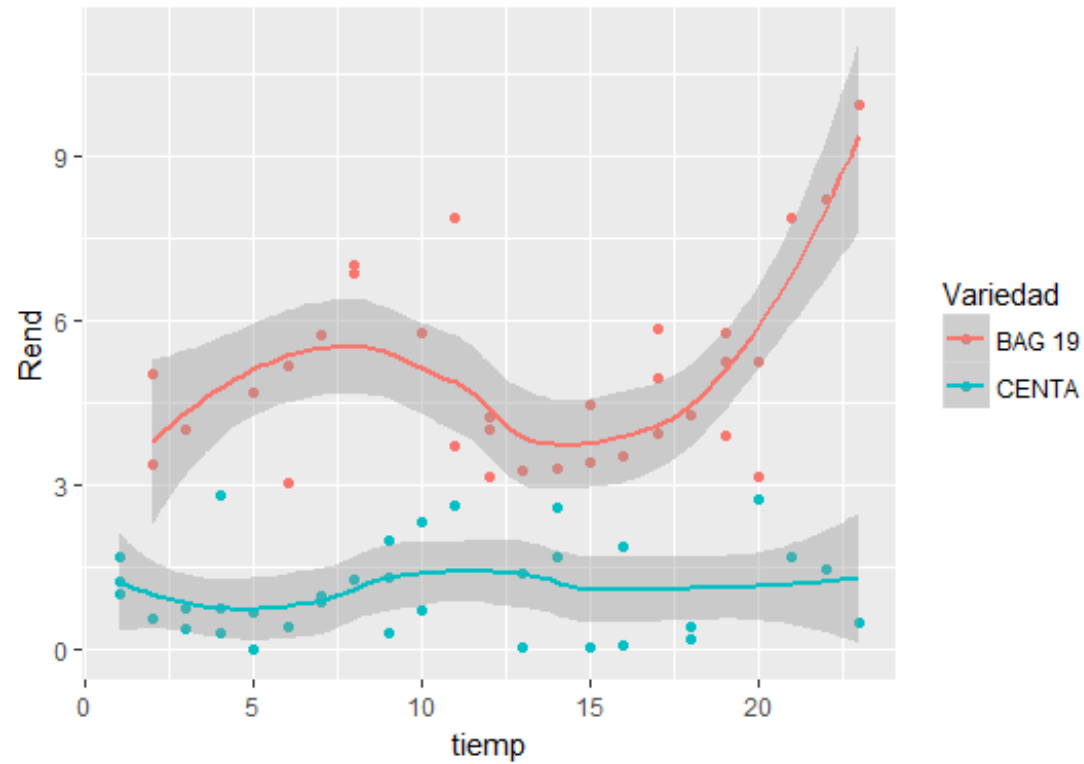
Comando básico de graficos

```
ggplot(trigo,aes(x=Rend))+geom_histogram(bins = 20,colour='green',fill='yellow')+ggtitle('Histograma de rendimiento de trigo')+ylab('rendimiento en kg/ha')
```



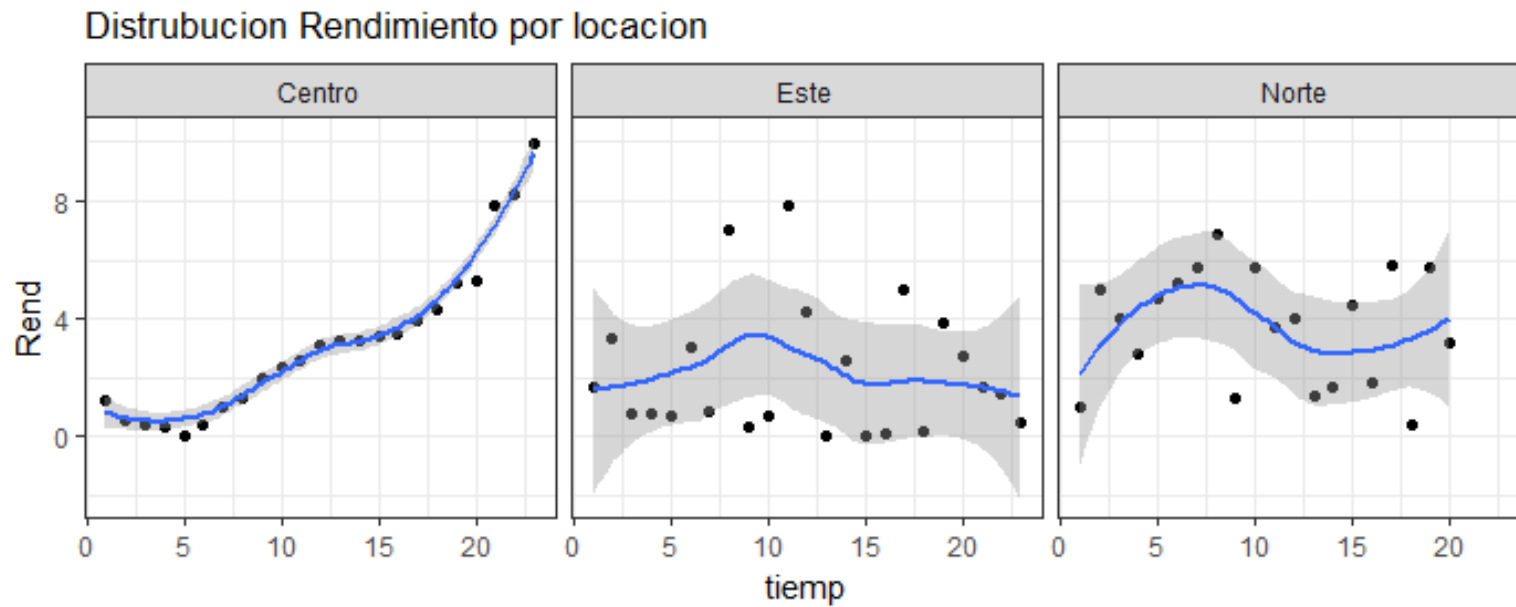
Colores y tendencia

```
ggplot(trigo,aes(x=tiemp,y =Rend ))+geom_point(aes(colour=Variedad))+geom_smooth(aes(colour=Variedad))
```



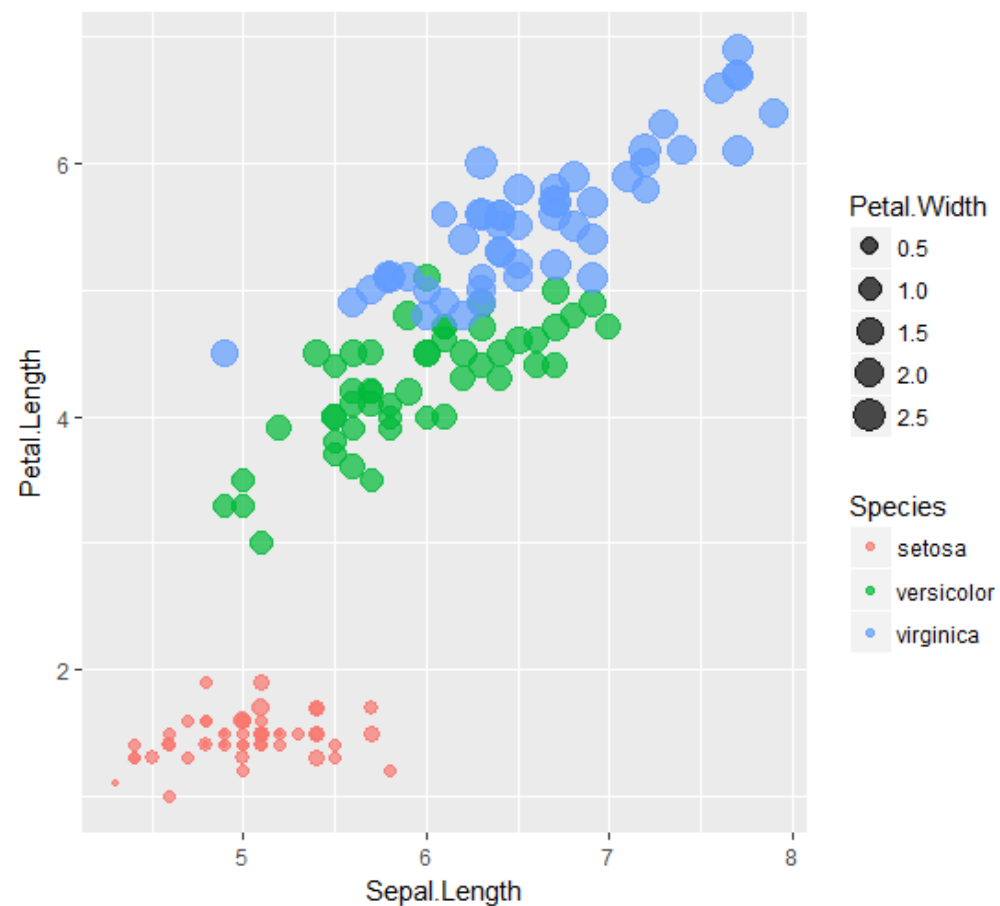
Facetas

```
ggplot(trigo,aes(x=tiemp,y =Rend ))+geom_point()+geom_smooth()+facet_grid(.~ubicacion) +theme_bw()+  
  ggtitle('Distrubucion Rendimiento por locacion')
```



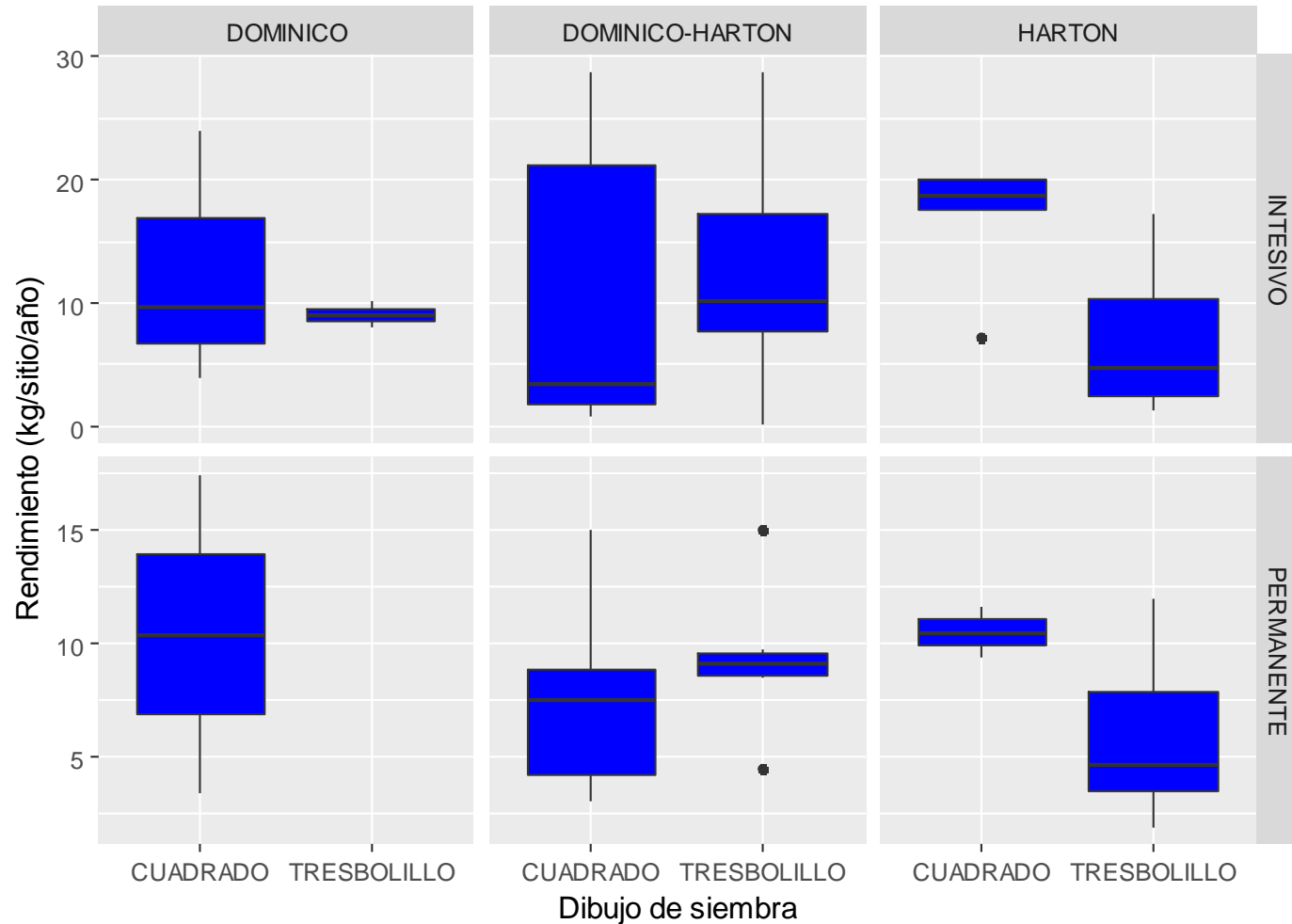
Tamaños, colores y transparencia

```
ggplot(trigo,aes(x=tiemp,y =Rend ))+geom_point()+geom_smooth()+facet_grid(.~ubicacion) +theme_bw()+  
  ggtitle('Distribucion Rendimiento por locacion')
```



Ejercicio

Rendimiento según dibujo de siembra, plátano Colombia 2013



1. Leer en R la bases de datos eventos_de_platano.csv.
2. Realizar una breve exploración de la base de datos.
3. Calcular la variable rendimiento dividiendo la producción sobre el número de árboles
4. Realizar el gráfico de la izquierda utilizando el paquete ggplot2

¡Gracias!



NOS ENORGULLECE
HABER CELEBRADO 50 AÑOS
DE INVESTIGACIÓN AGRÍCOLA
PARA EL DESARROLLO

Centro Internacional de Agricultura Tropical - CIAT

Sede Principal y Oficina Regional
para Suramérica y el Caribe

+57 2 445 0000

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

✉ ciat@cgiar.org

🌐 ciat.cgiar.org



El CIAT es un Centro de Investigación de CGIAR