

# Introducción a la visualización de datos

Arturo Pérez

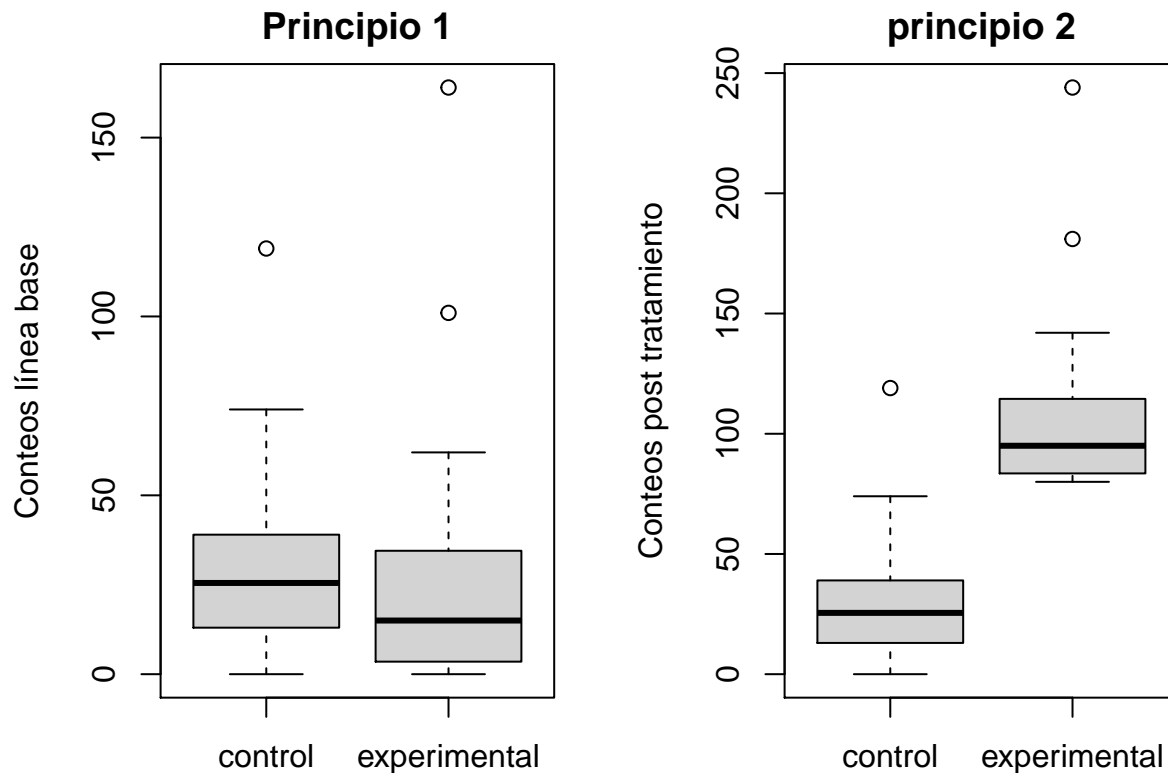
## 1. Principios de las gráficas de análisis

### Mostrar comparaciones.

La evidencia de una hipótesis siempre es relativa si no tenemos un punto de comparación. Es decir, siempre hay que mostrar un contraste de condiciones (p.e Control y experimental)

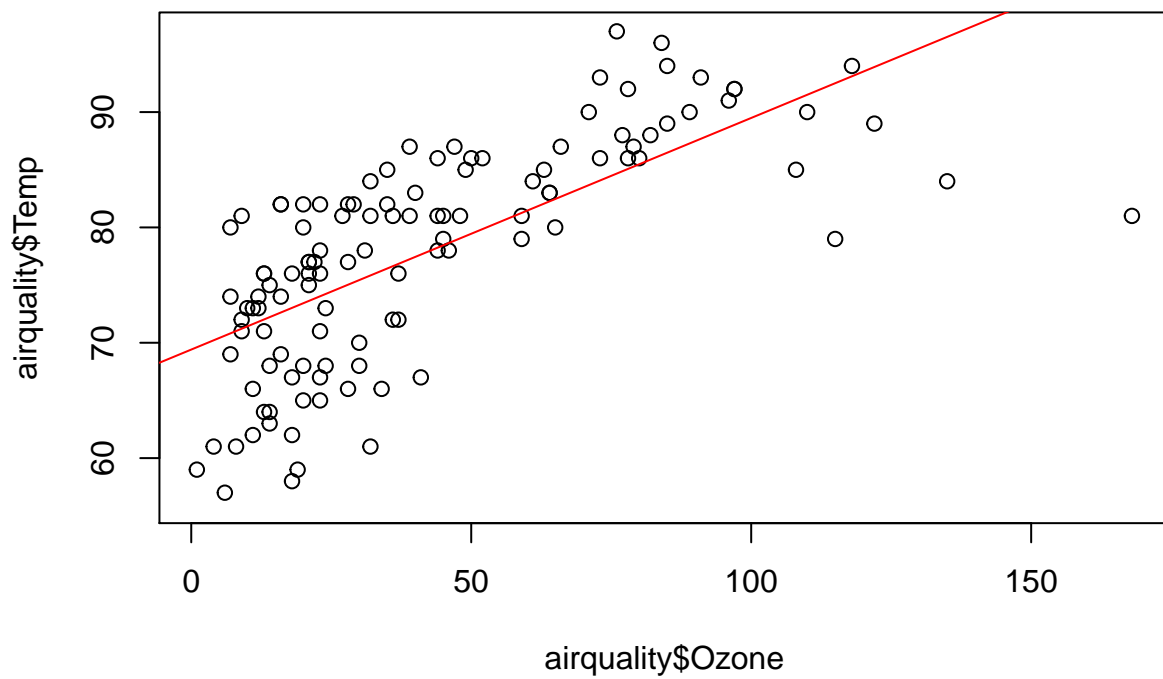
### Mostrar causalidad, mecanismos, estructura sistemática.

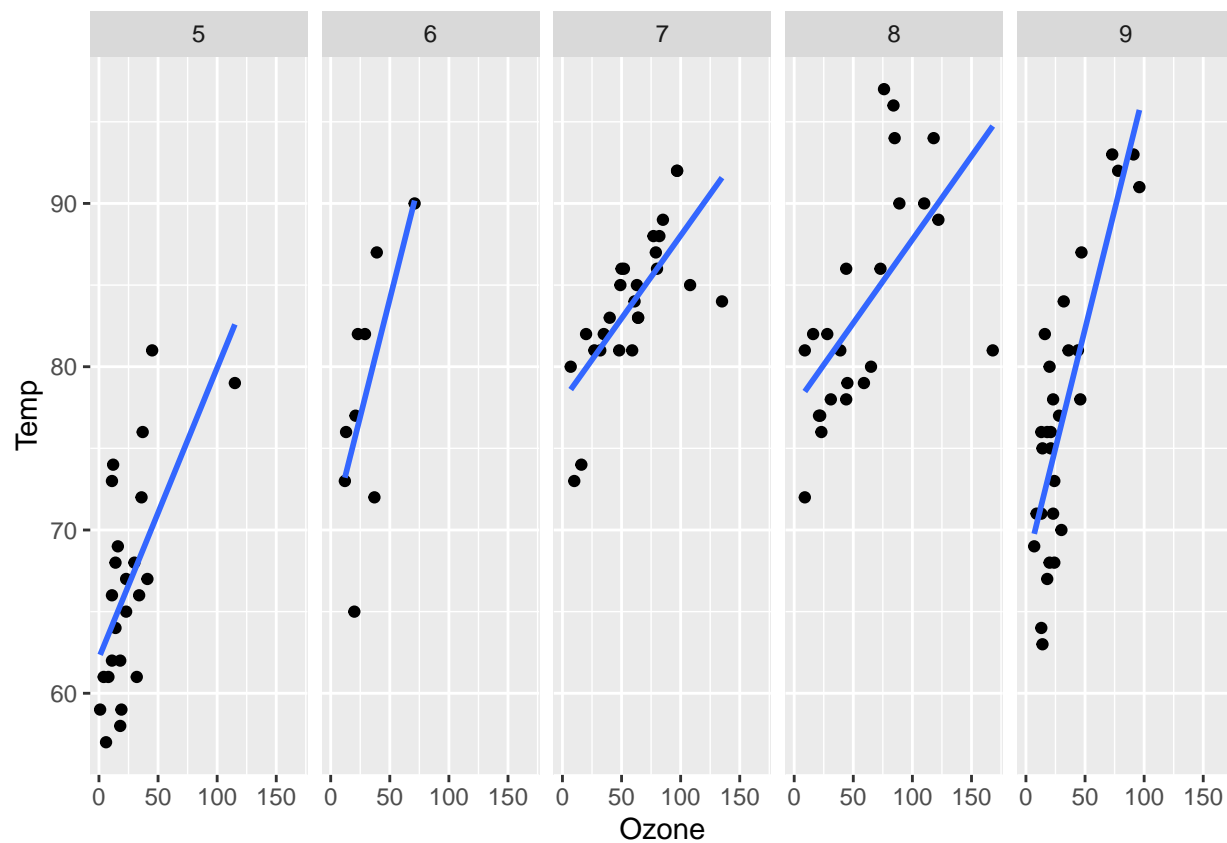
Es básicamente cómo creo que se comporta el fenómeno



### Mostrar datos multivariados

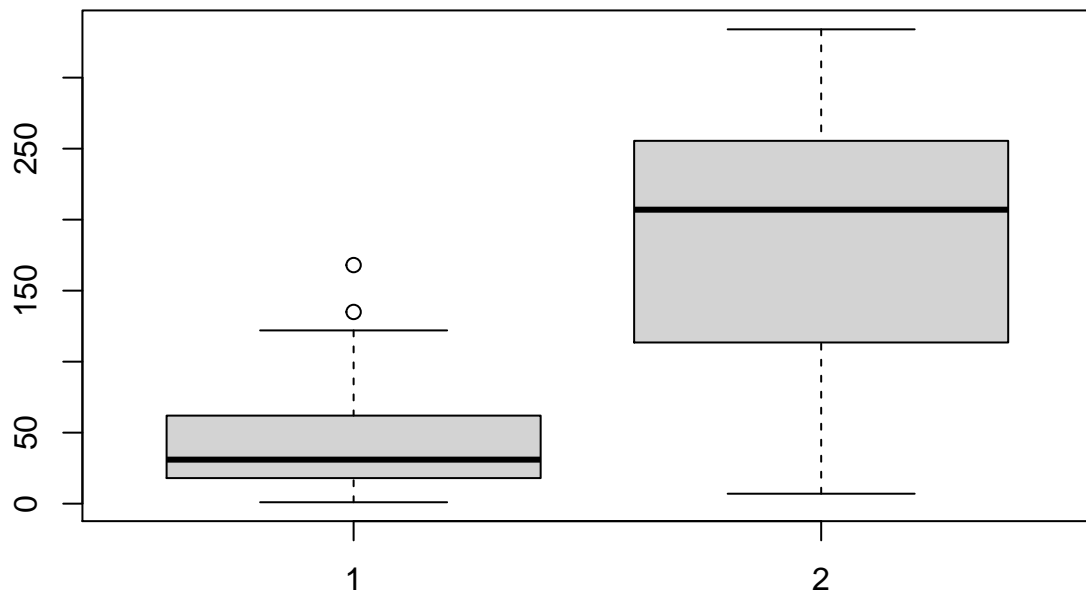
Mostrar más de dos variables. El mundo real es multivariado.



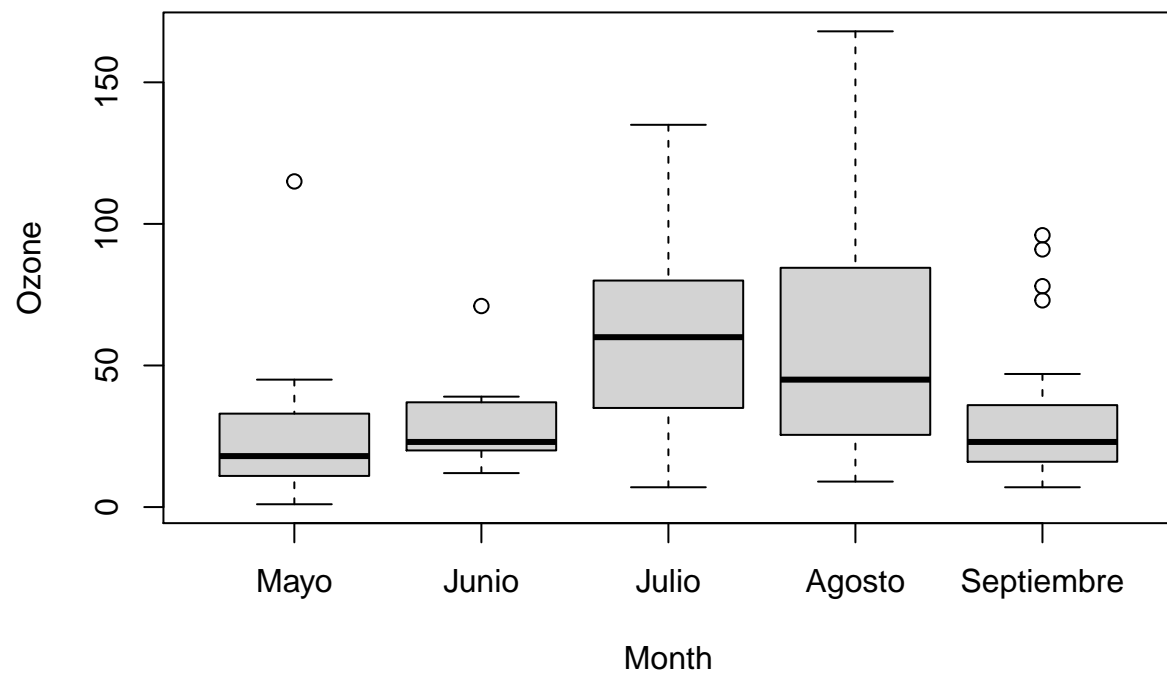


## 2. Diagramas de cajas y bigotes

```
boxplot(airquality$Ozone, airquality$Solar.R)
```

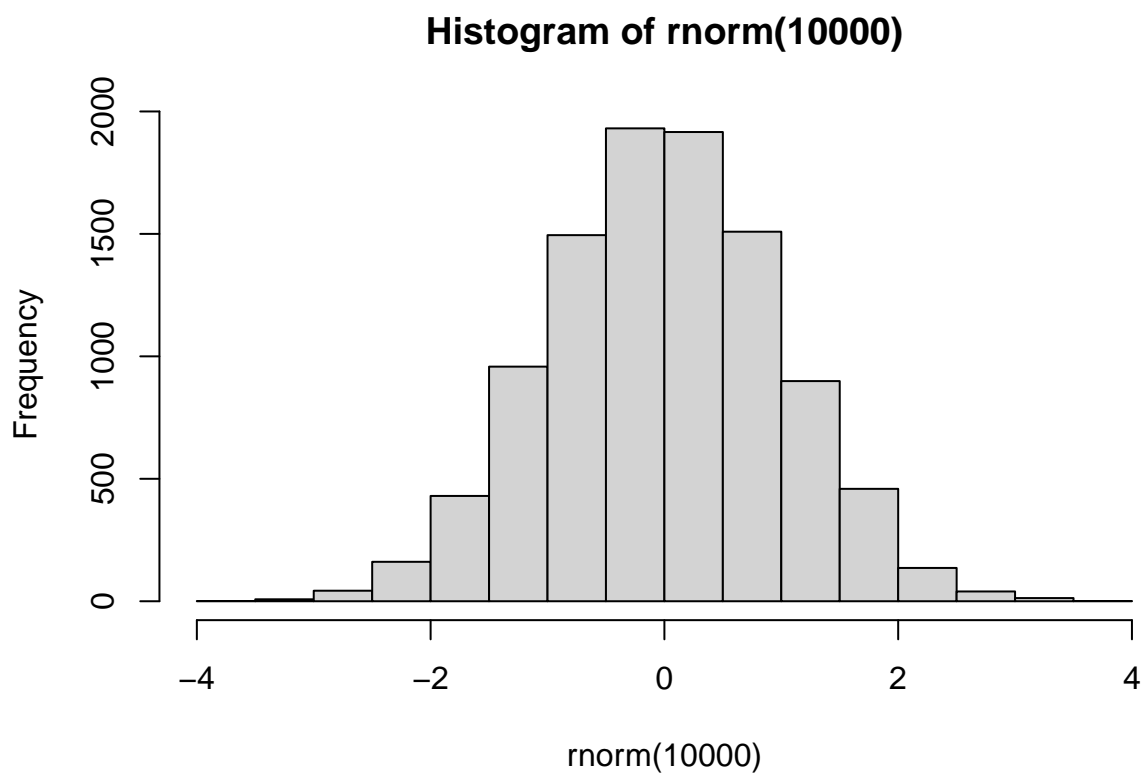


```
airquality$Month <- factor(airquality$Month,  
                           labels = c("Mayo", "Junio", "Julio", "Agosto", "Septiembre"))  
boxplot(Ozone~Month, data = airquality)
```



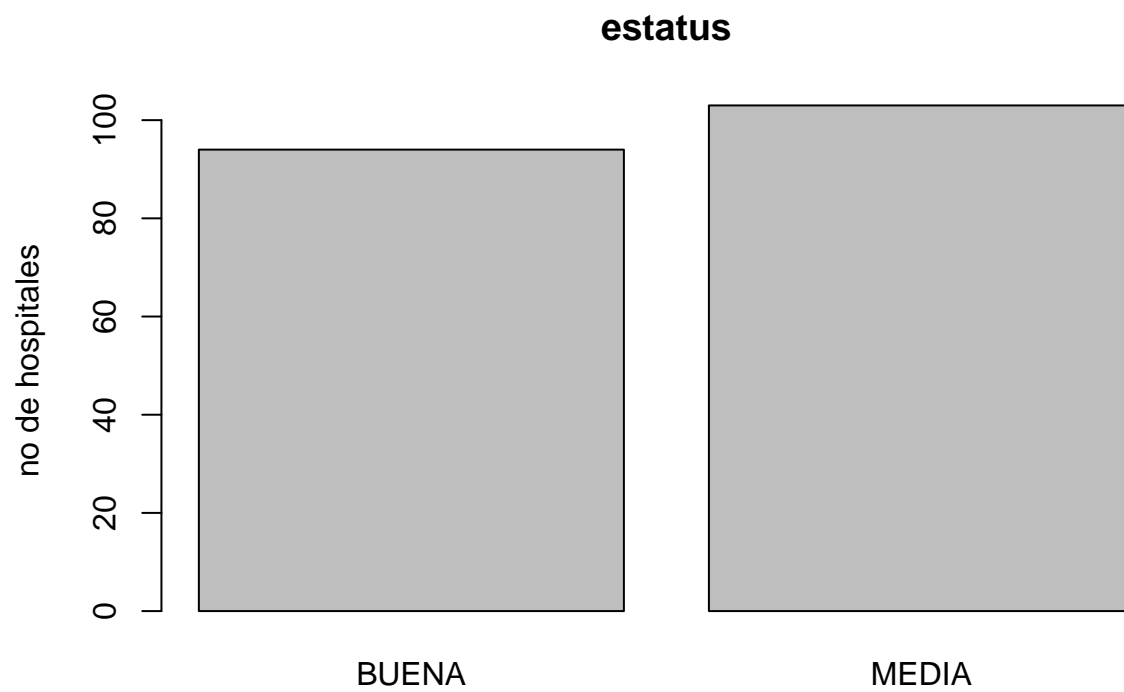
### 3. Histograma

```
hist(rnorm(10000))
```



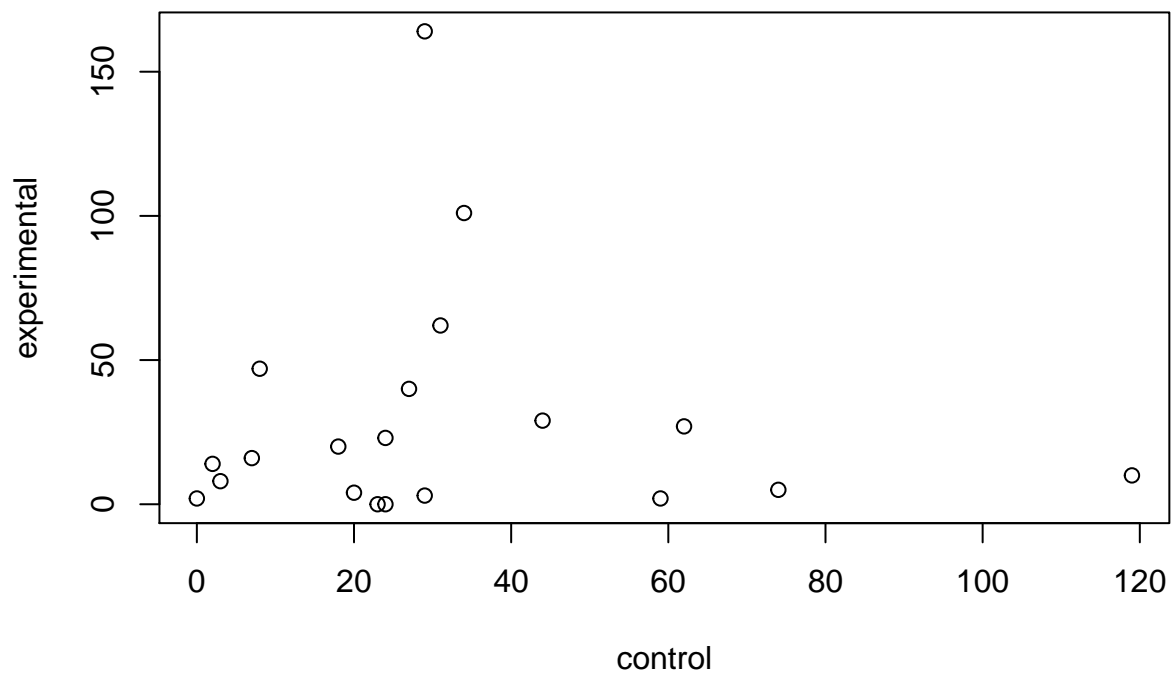
#### 4. Gráfica de barras

```
estatus <- read.csv("./covid/capacidadcovid19CDMX.csv")
barplot(table(estatus$Estatus_capacidad_hospitalaria[estatus$Estatus_capacidad_hospitalaria=="BUENA",
                                                       "MEDIA"]),
        , main = "estatus", ylab = "no de hospitales")
```



## 5. Diagramas de dispersión

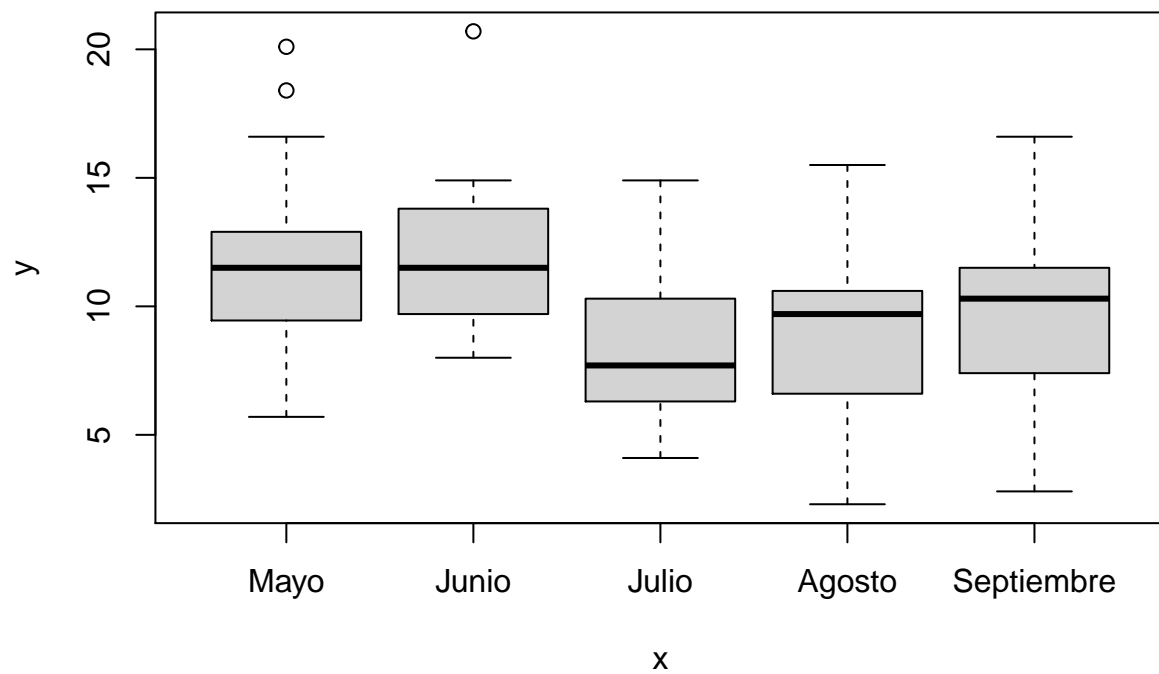
```
plot(control, experimental)
```



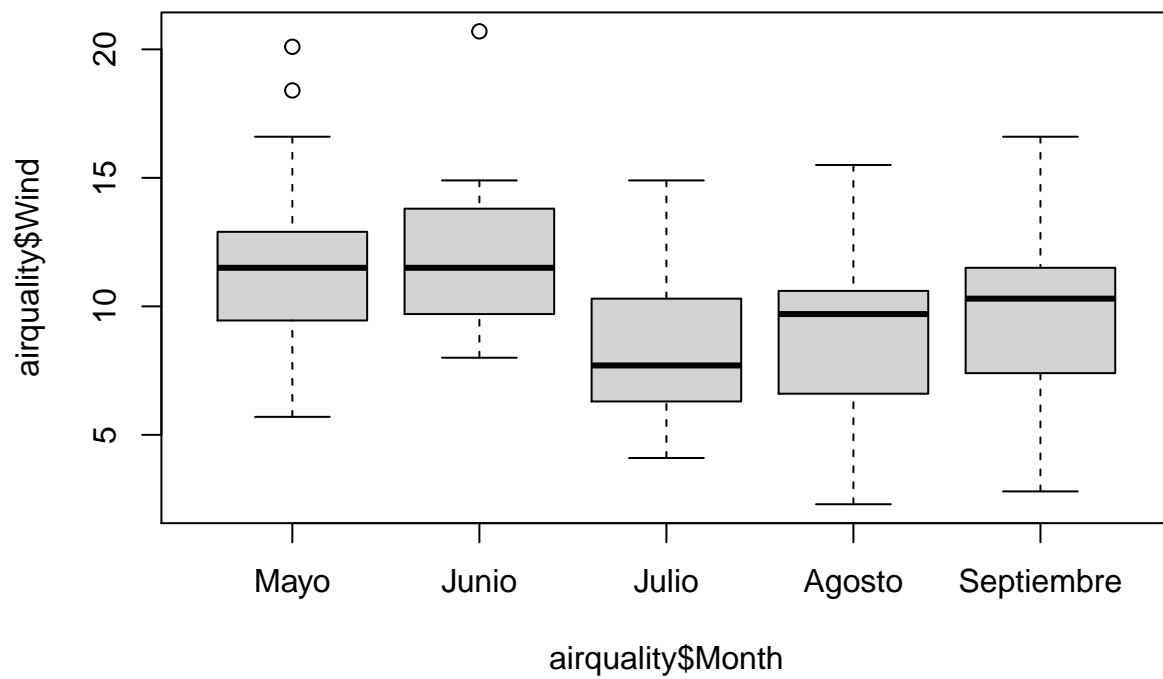
De acuerdo al tipo de variable que vayamos a graficar podemos obtener un diagrama de dispersión o un boxplot (cuando hay variables categóricas).

```
plot(airquality$Month,airquality$Wind)
```



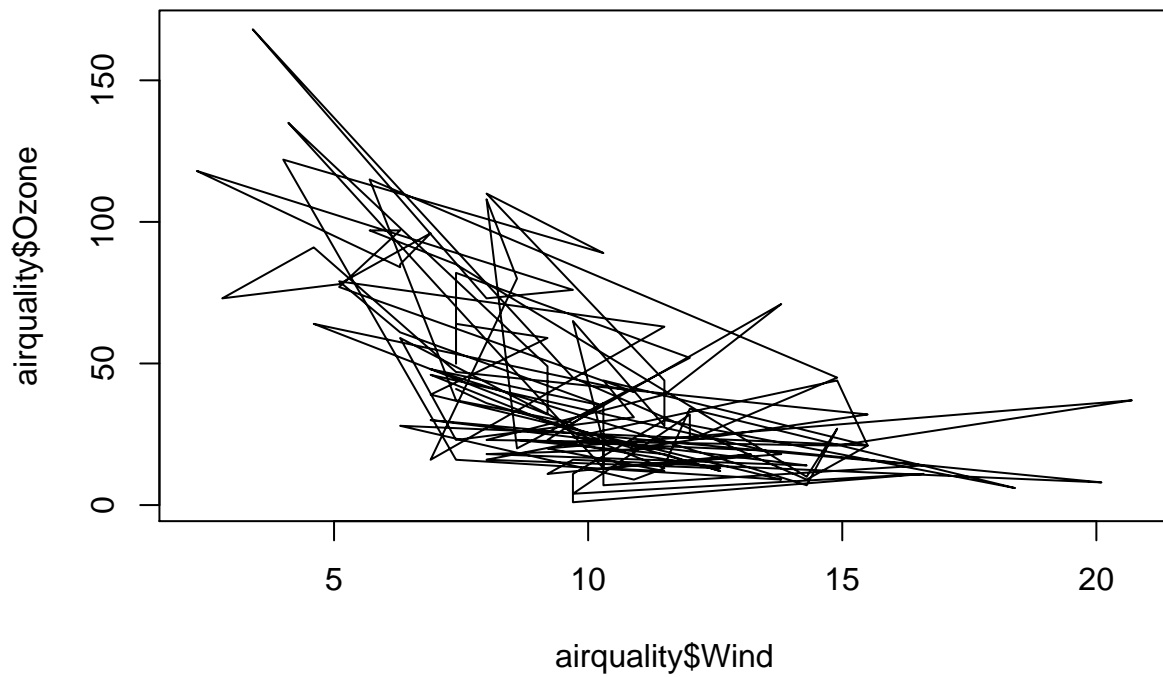


```
plot(airquality$Wind~airquality$Month)
```

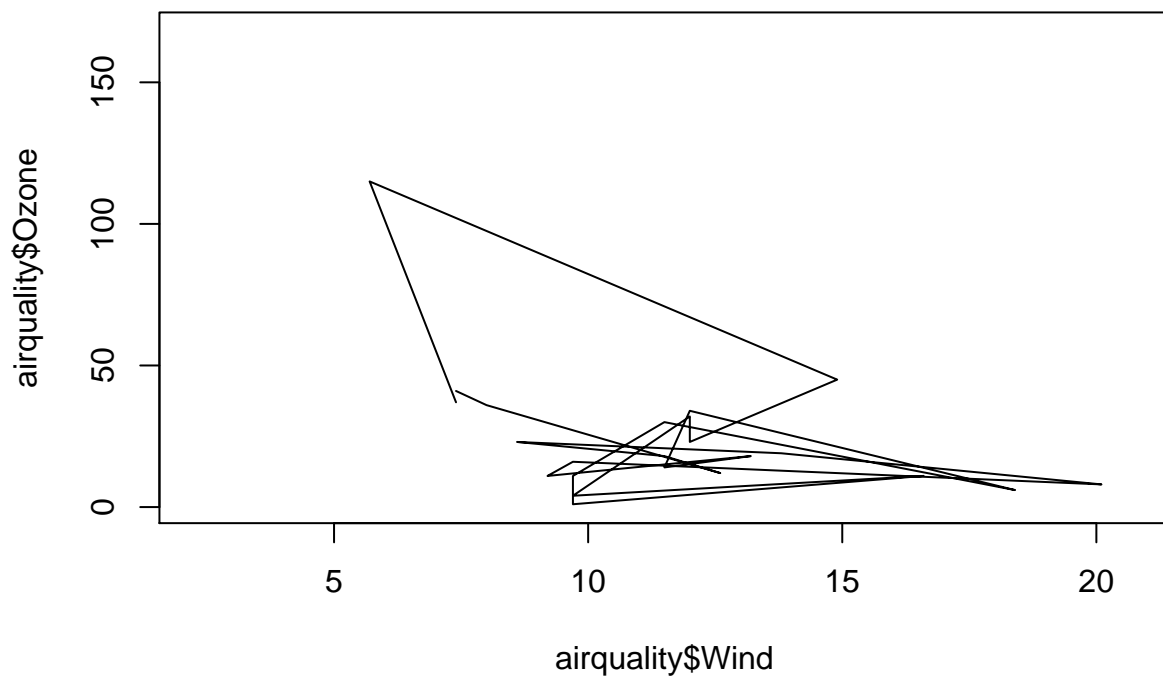


## 6. Funciones de graficado.

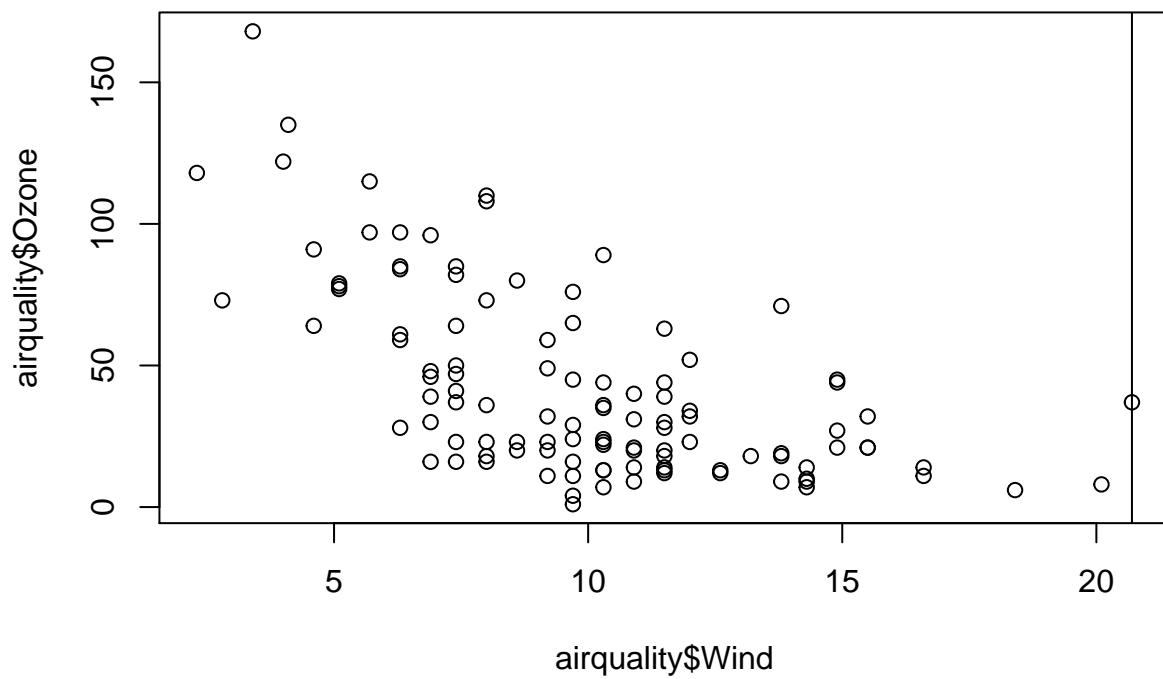
```
#lines  
plot(airquality$Wind,airquality$Ozone, type = "n")  
lines(airquality$Wind,airquality$Ozone)
```



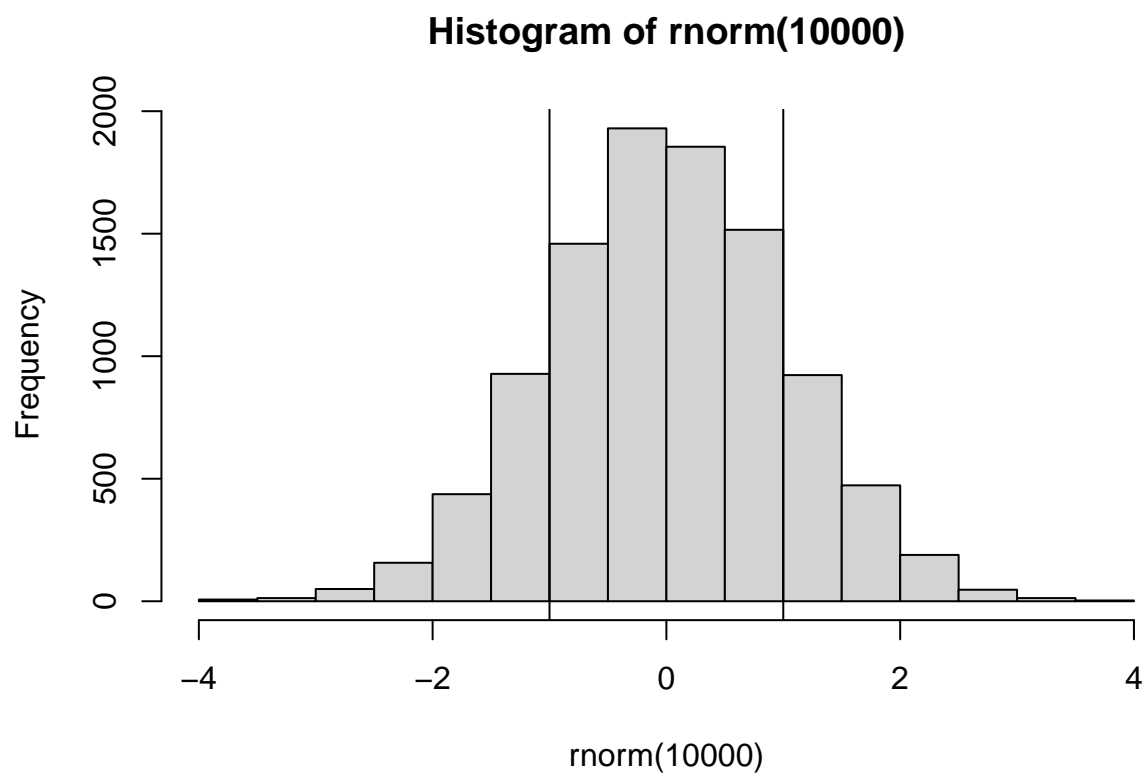
```
##cuando quiero líneas de un solo subset  
plot(airquality$Wind,airquality$Ozone, type = "n")  
with(subset(airquality, Month=="Mayo"),lines(Wind, Ozone))
```



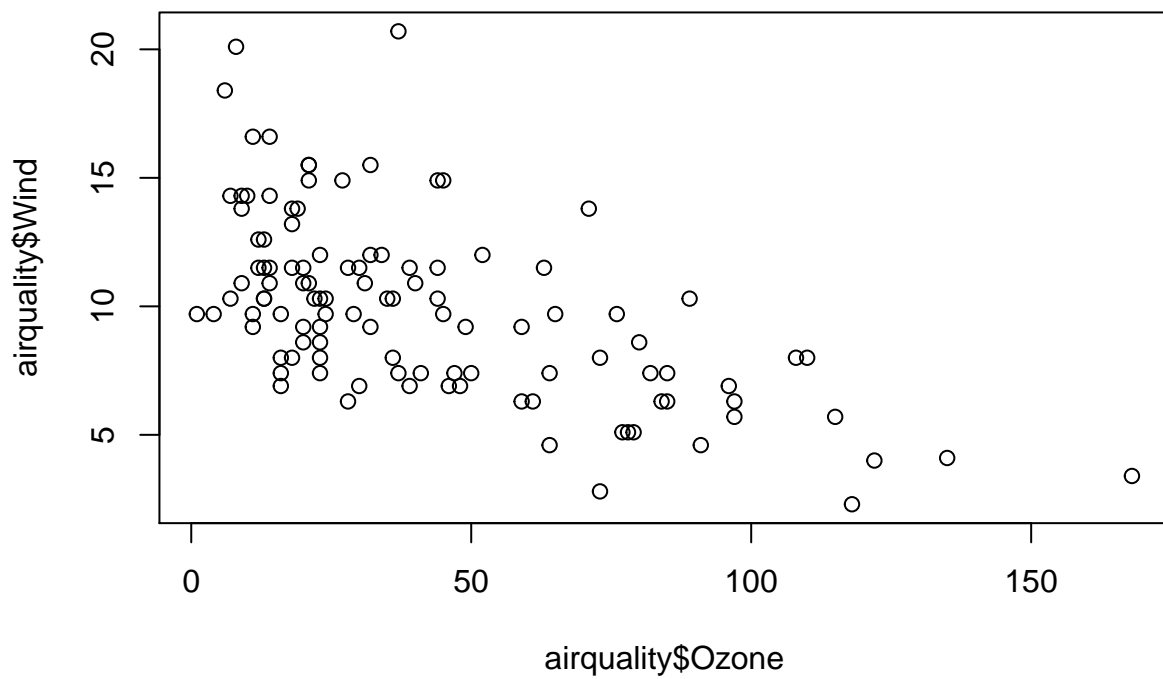
```
#abline  
plot(airquality$Wind,airquality$Ozone)  
abline(v=max(airquality$Wind))
```



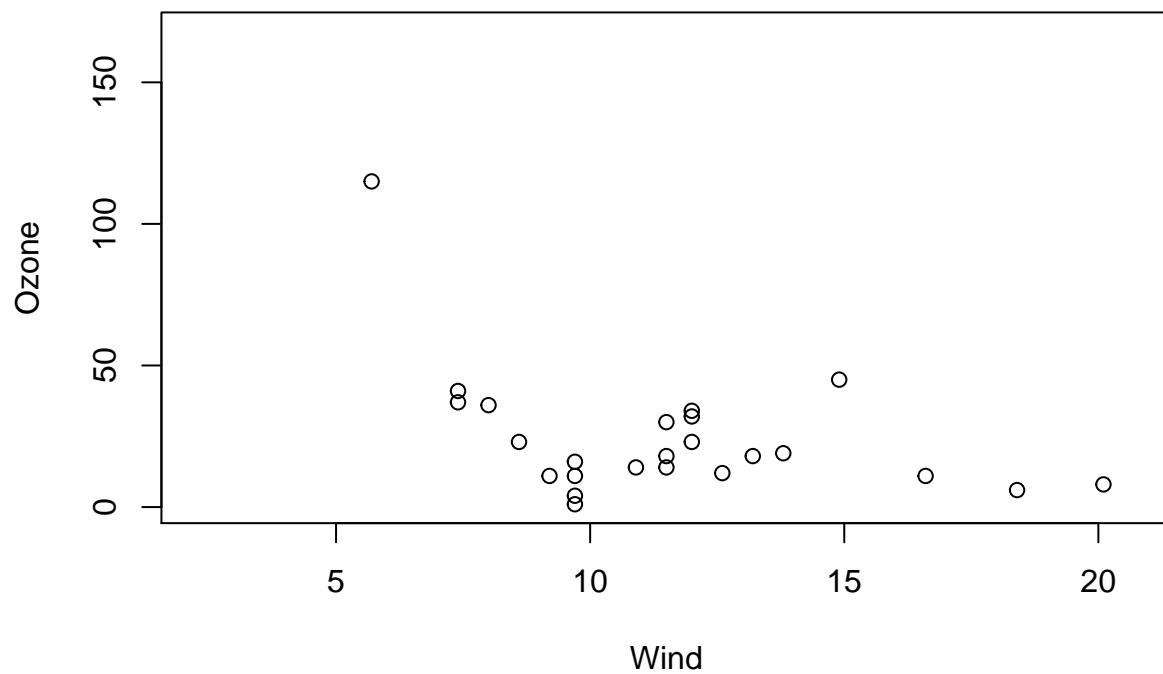
```
hist(rnorm(10000))  
abline(v=c(-1,1))
```



```
#points  
plot(airquality$Ozone, airquality$Wind, type = "n")  
points(airquality$Ozone, airquality$Wind)
```

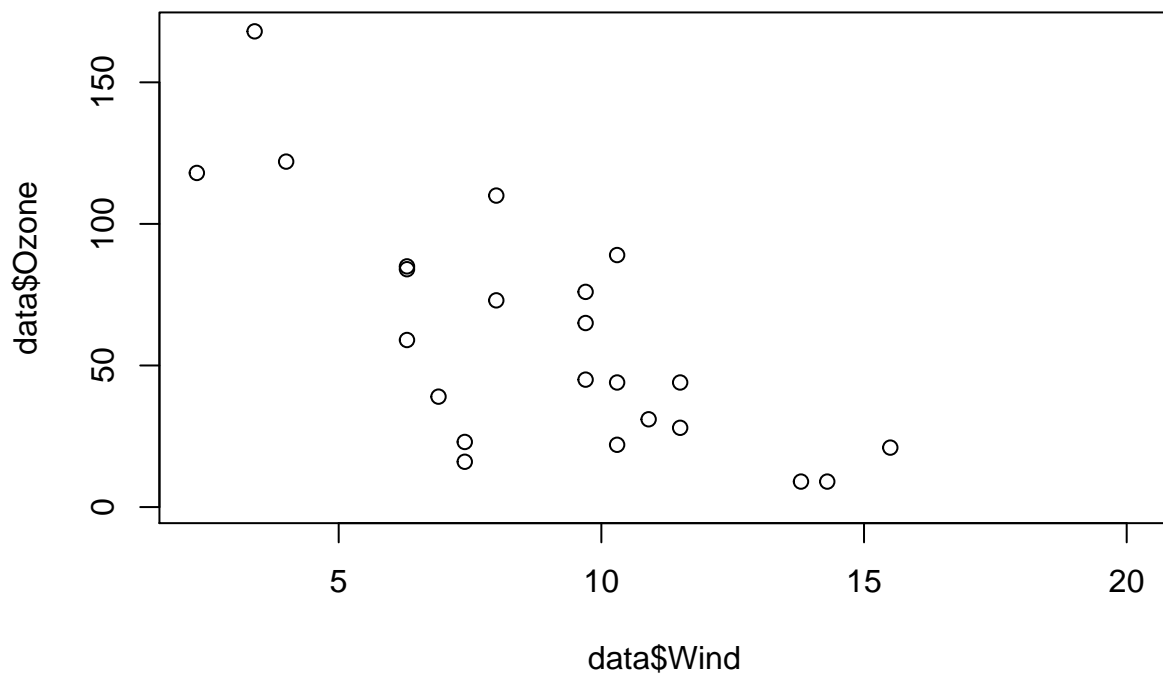


```
##sólo un grupo de puntos  
with(airquality, plot(Wind, Ozone, type="n"))  
with(subset(airquality, Month=="Mayo"), points(Wind, Ozone) )
```

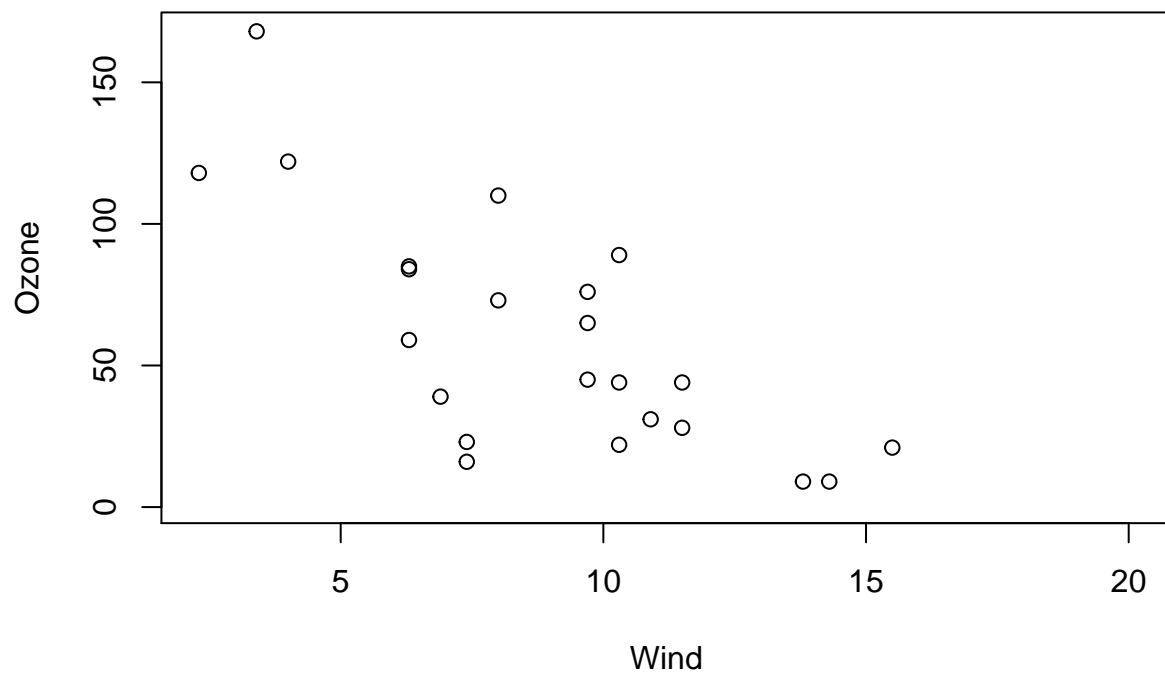


```
##meses específicos
data <- airquality[airquality$Month %in% c("Mayo", "Agosto"),]
plot(data$Wind, data$Ozone, type = "n")
#puntos de agosto
points(data[data$Month=="Agosto", "Wind"],data[data$Month=="Agosto", "Ozone"] )
```

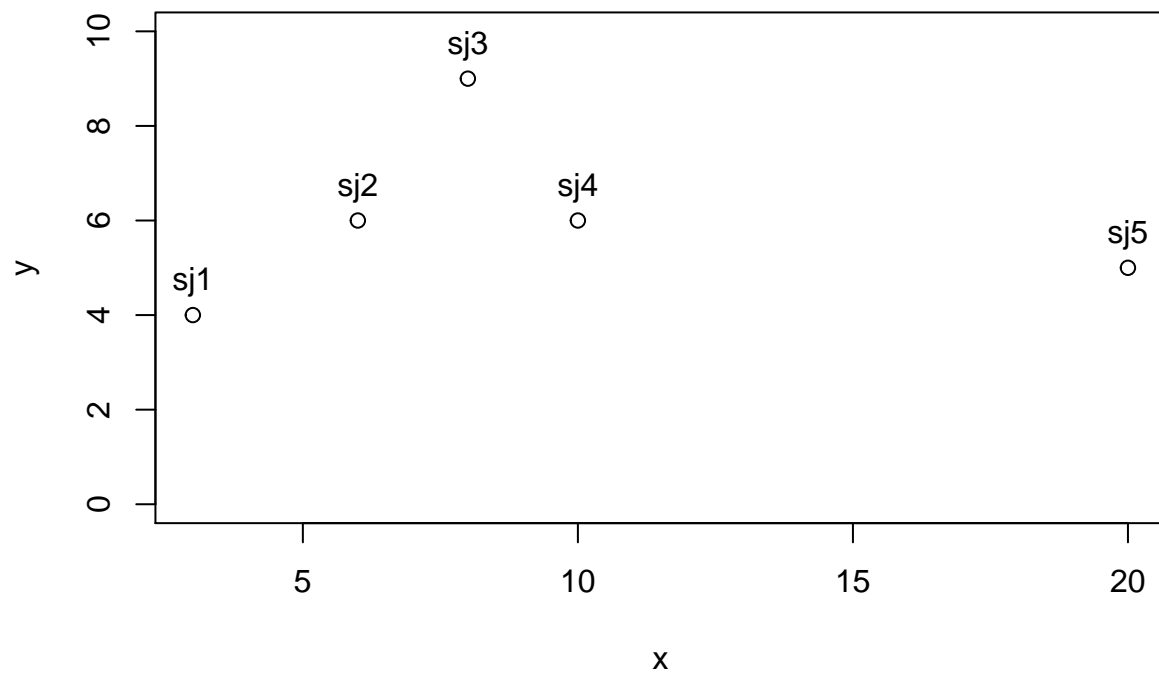




```
##EQUIVALENTE A LA GRAFICA ANTERIOR
with(airquality[airquality$Month %in% c("Mayo", "Agosto"),],
      plot(Wind, Ozone, type = "n"))
#puntos de agosto
with(subset(airquality, Month == "Agosto"), points(Wind, Ozone))
```

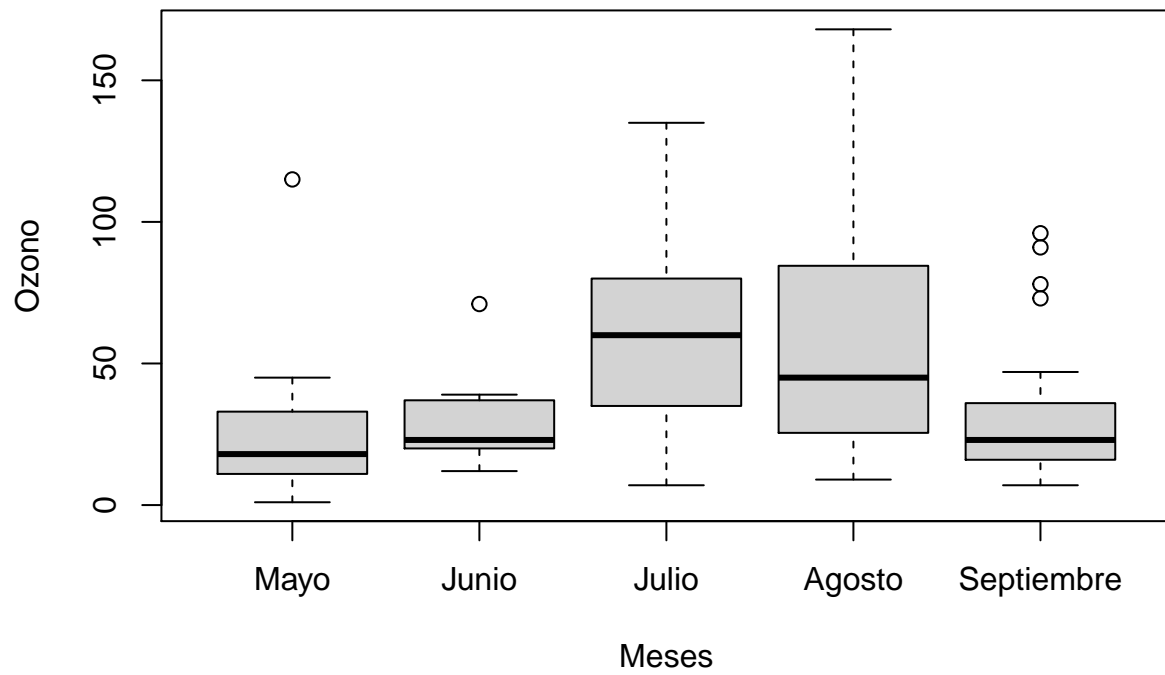


```
#text
plot(c(3,6,8,10,20), c(4,6,9,6,5), xlab = "x", ylab = "y", ylim = c(0,10))
text(c(3,6,8,10,20), c(4,6,9,6,5)+.7, labels = c("sj1", "sj2", "sj3", "sj4", "sj5"))
```

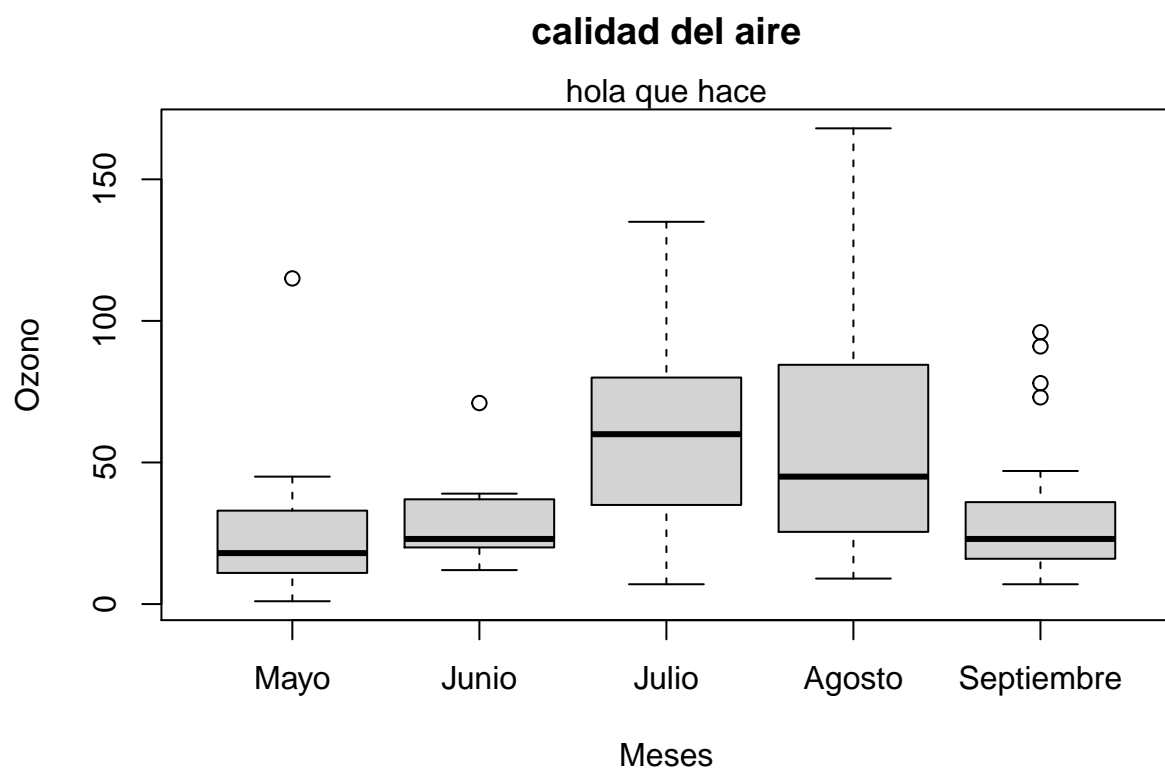


```
##title  
boxplot(Ozone~Month, data = airquality, ylab = "", xlab = "")  
title(main = "Calidad del aire", xlab = "Meses", ylab = "Ozono")
```

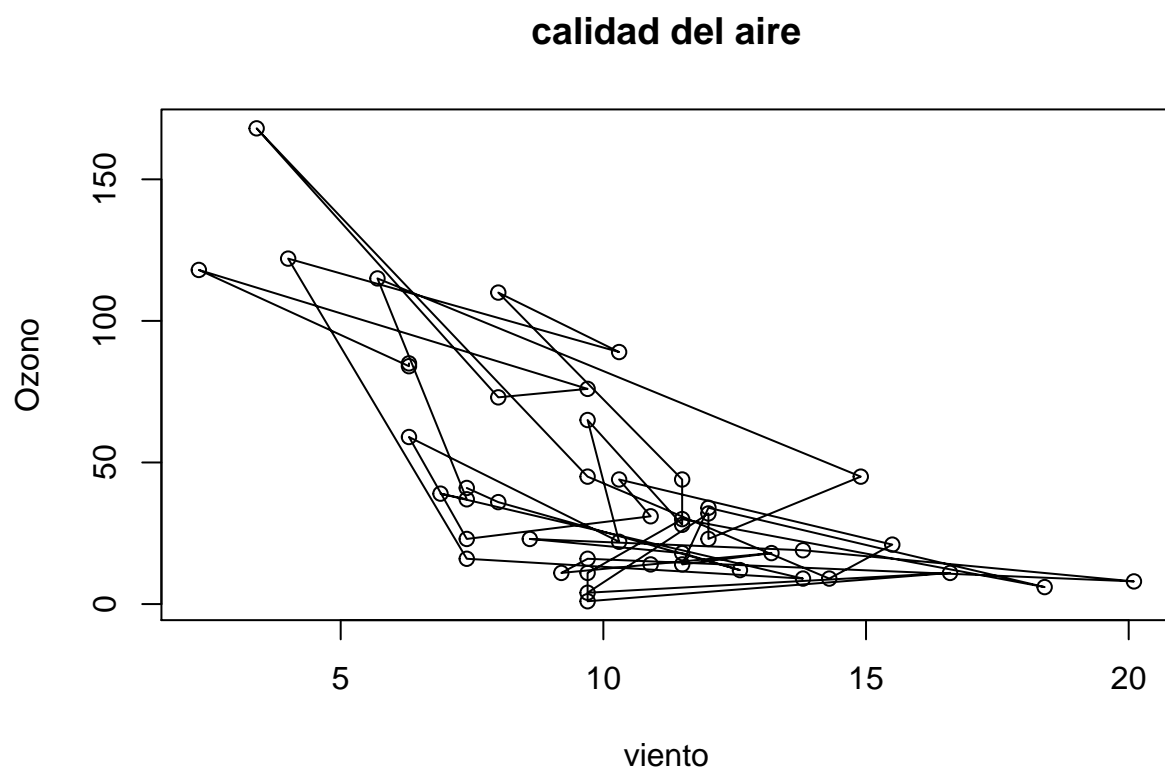
## Calidad del aire



```
##mtext
boxplot(Ozone~Month, data = airquality, main="calidad del aire", xlab = "Meses", ylab = "Ozono")
mtext("hola que hace", 3)
```

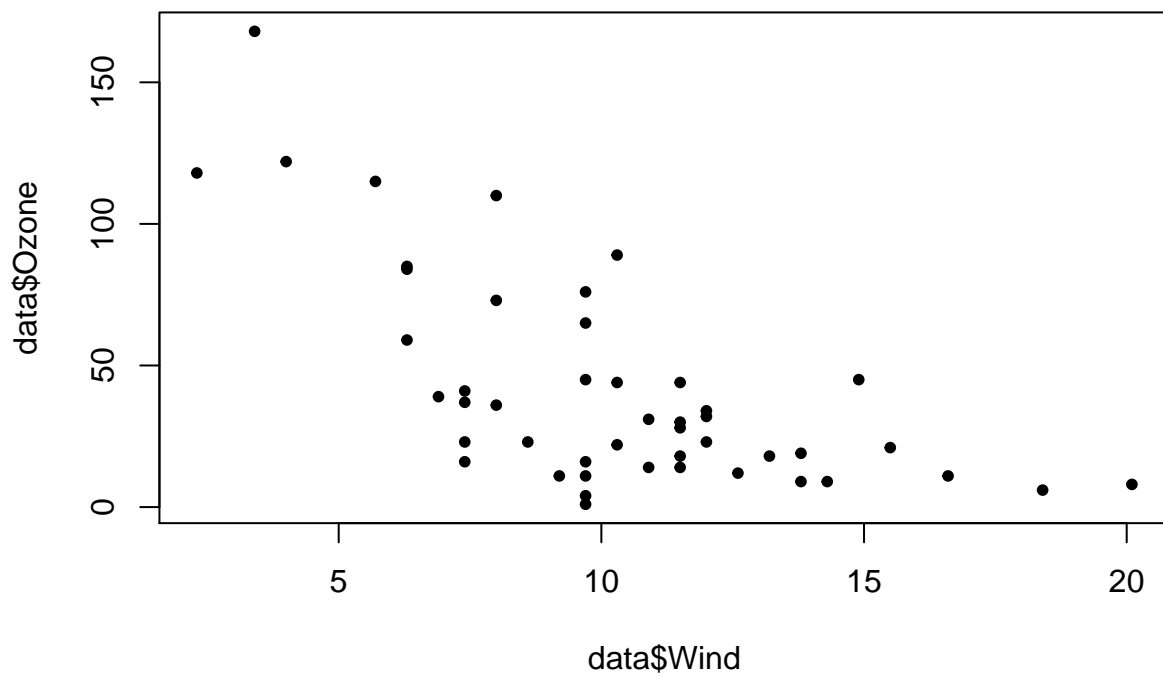


```
#Algunas funciones se pueden aplicar directamente en
with(airquality[airquality$Month %in% c("Mayo", "Agosto"),],
      plot(Wind, Ozone, main = "calidad del aire", xlab = "viento", ylab = "Ozono", type = "o" ))
```



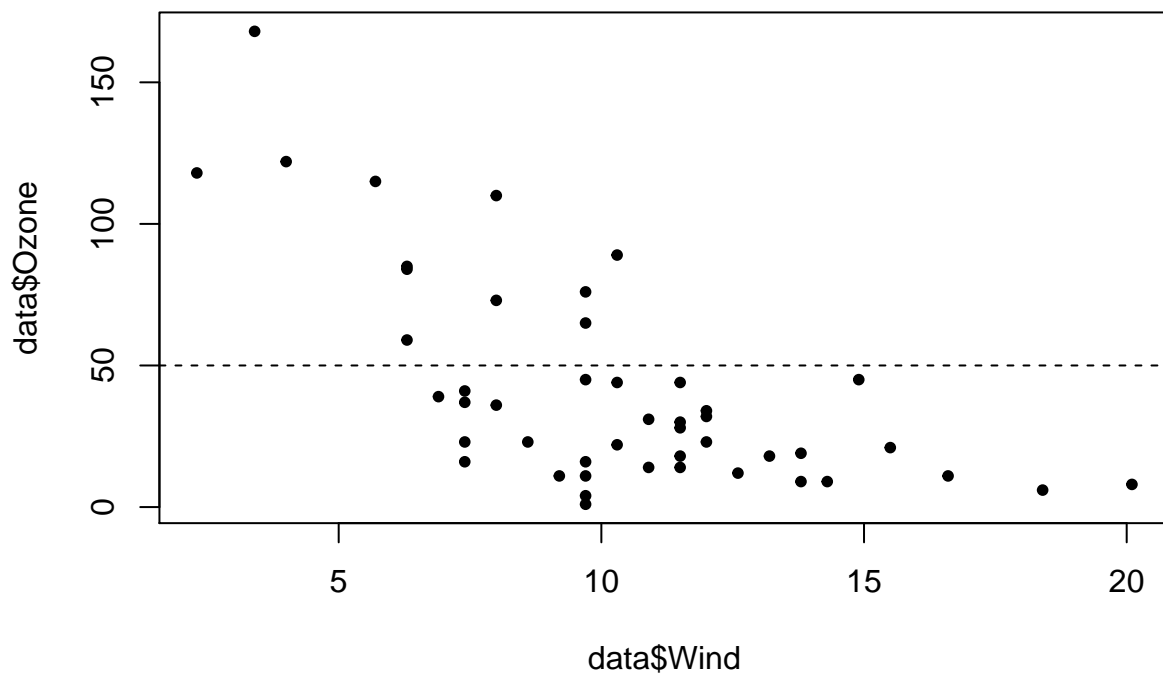
### Características especiales dentro de una gráfica **pch**

```
plot(data$Wind, data$Ozone, pch=20)
```



lty

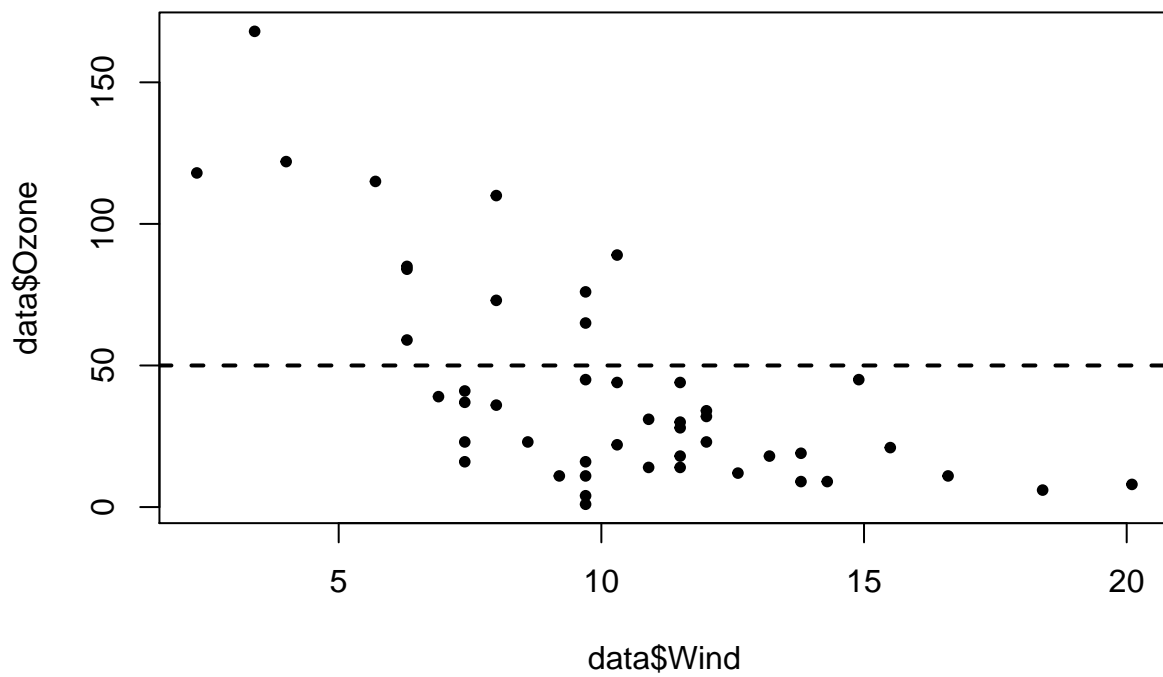
```
plot(data$Wind, data$Ozone, pch=20)  
abline(h=50, lty=2)
```



lwd

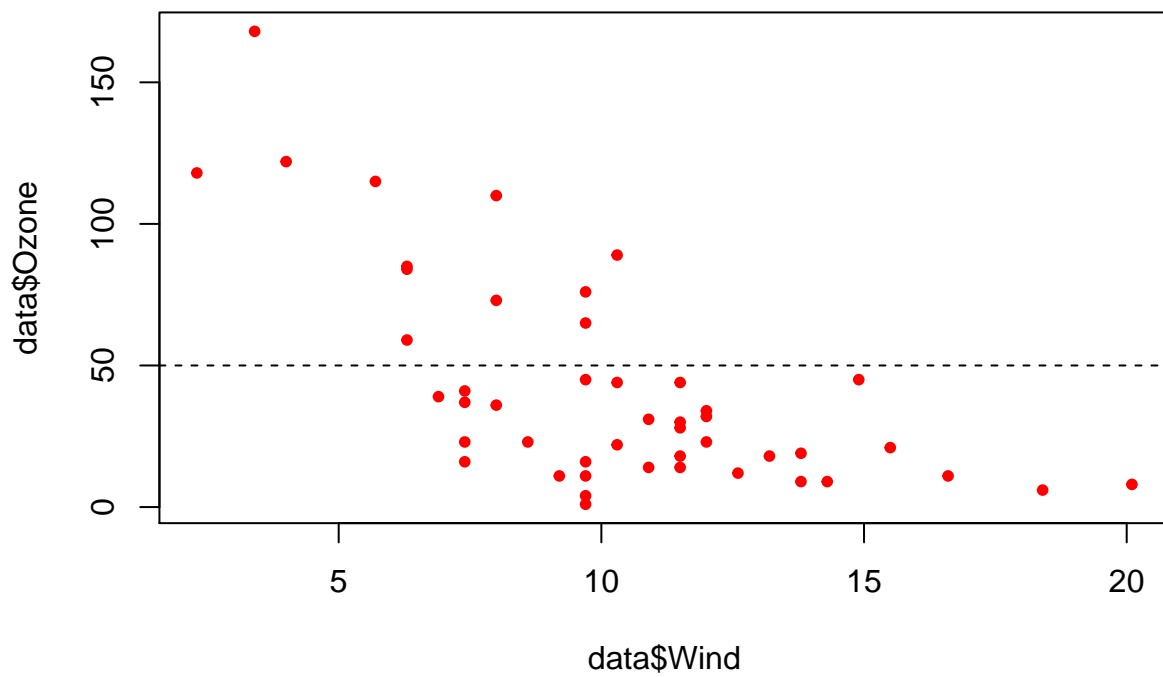
```
plot(data$Wind, data$Ozone, pch=20)
abline(h=50, lty=2, lwd=2)
```



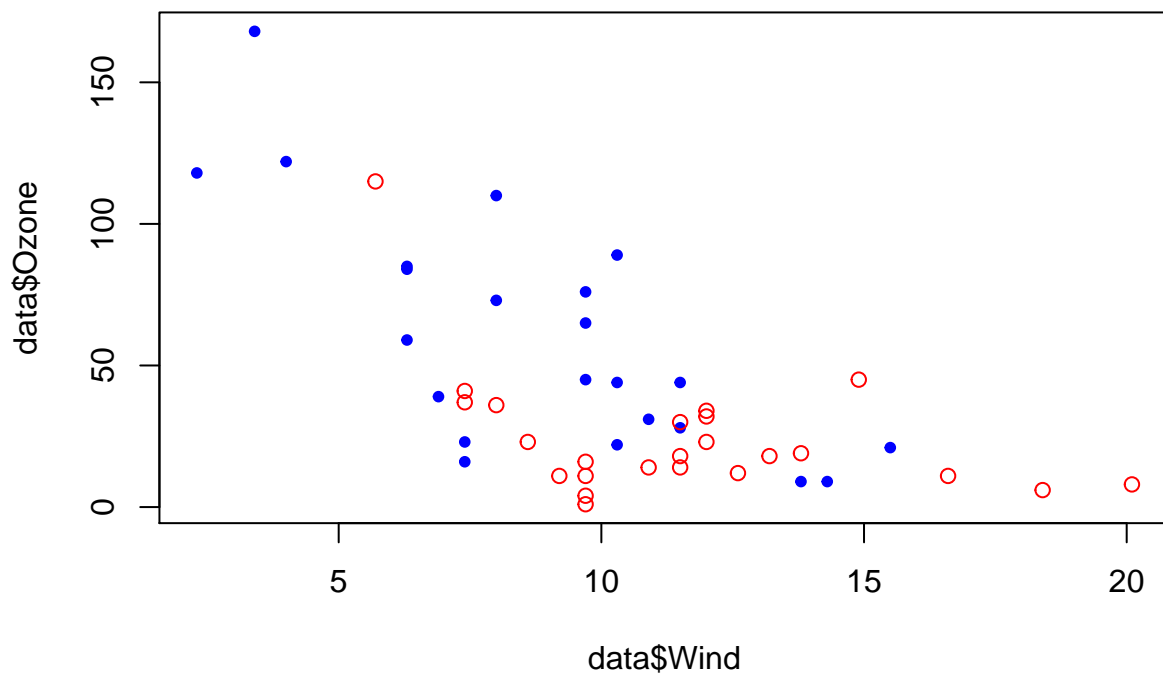


col

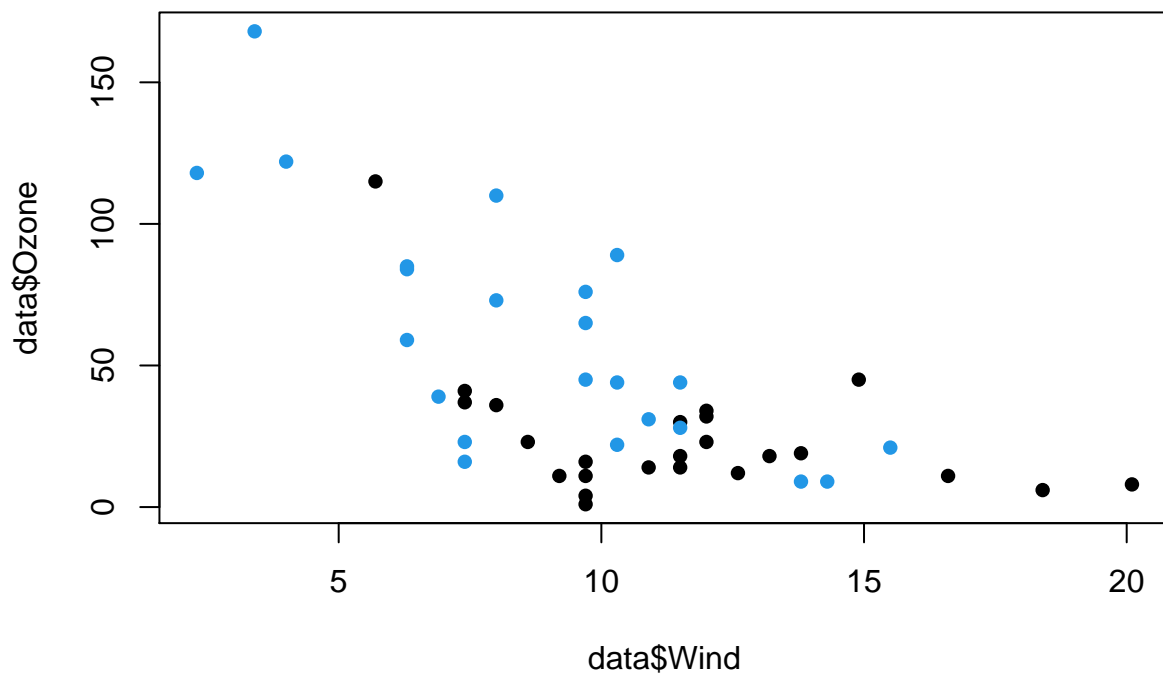
```
plot(data$Wind, data$Ozone, pch=20, col="red")  
abline(h=50, lty=2, lwd=1)
```



```
plot(data$Wind, data$Ozone, type="n")
points(data[data$Month=="Agosto", "Wind"],
       data[data$Month=="Agosto", "Ozone"], col="blue", pch=20)
points(data[data$Month=="Mayo", "Wind"],
       data[data$Month=="Mayo", "Ozone"], col="red", pch=1)
```

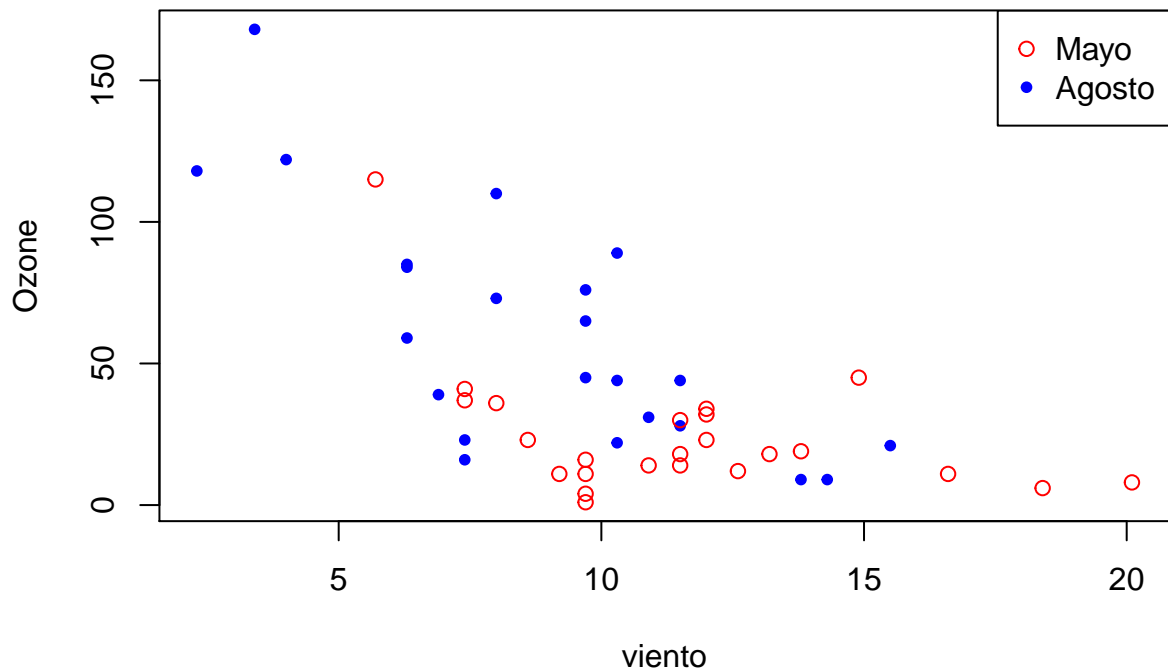


```
plot(data$Wind, data$Ozone, pch=16, col=data$Month)
```



legend

```
plot(data$Wind, data$Ozone, type="n", xlab = "viento", ylab = "Ozone")
points(data[data$Month=="Agosto", "Wind"],
       data[data$Month=="Agosto", "Ozone"], col="blue", pch=20)
points(data[data$Month=="Mayo", "Wind"],
       data[data$Month=="Mayo", "Ozone"], col="red", pch=1)
legend("topright", legend = c("Mayo", "Agosto"), pch = c(1,20), col = c("red", "blue"))
```



Guardando una gráfica

```
png("./plots/grafica1")
plot(data$Wind, data$Ozone, type="n", xlab = "viento", ylab = "Ozone")
points(data[data$Month=="Agosto", "Wind"],
       data[data$Month=="Agosto", "Ozone"], col="blue", pch=20)
points(data[data$Month=="Mayo", "Wind"],
       data[data$Month=="Mayo", "Ozone"], col="red", pch=1)
legend("topright", legend = c("Mayo", "Agosto"), pch = c(1,20), col = c("red", "blue"))
```