

Datos Hidrológicos; Ejercicio explorativo.

Debemos importar los datos a R con el siguiente comando:

```
inp <- read.csv("FDC.csv", na.strings="")
```

Con la función **head** se puede visualizar el encabezado del documento:

```
head(inp)
```

```
##      Tiempo Pandora.mm..637km2. Banano.mm..90km2.
## 1 6/22/1973                7.46                21.02
## 2 6/23/1973                6.17                14.78
## 3 6/24/1973                6.32                27.65
## 4 6/25/1973                7.80                19.68
## 5 6/26/1973                7.72                15.17
## 6 6/27/1973                7.83                23.14
```

Para ver la longitud de las series o ver cuantas filas y columnas hay en el archivo, usaremos la función **dim**:

```
dim(inp)
```

```
## [1] 3845    3
```

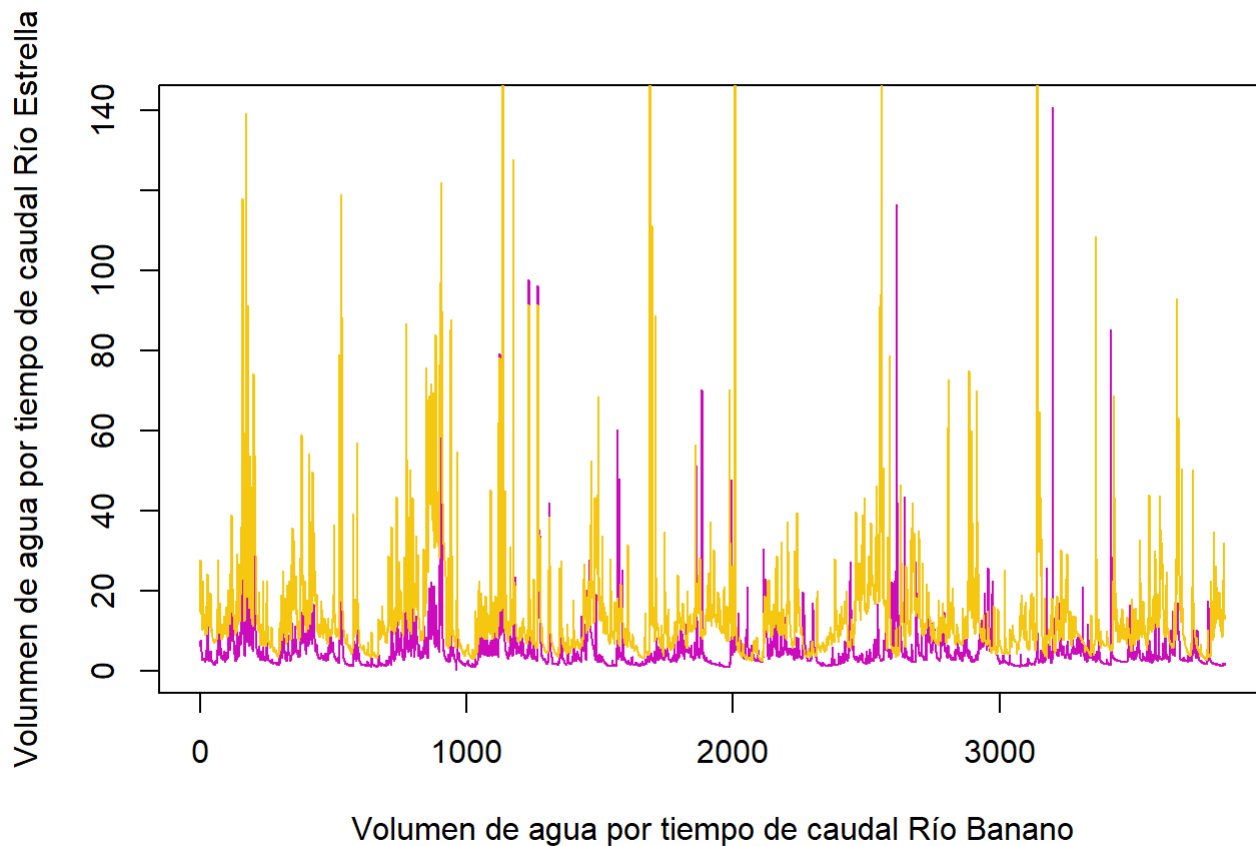
Para verificar que todas las filas son casos completos se utiliza el siguiente comando:

```
inp [!complete.cases(inp),]
```

Grafico del Volumen de agua por tiempo de los caudales de los ríos Banano y Estrella:

Utilizamos el comando **plot** para crear graficos y **xlab**, **ylab** para nombrar los ejes:

```
plot(inp[,2], type = "l", col= 6, xlab= "Volumen de agua por tiempo de caudal Río Banano", ylab=
"Volunmen de agua por tiempo de caudal Río Estrella")
lines(inp[,3], col= 7)
```



Para caracterizar con una estadística descriptiva básica y visualizar los promedios de los caudales diarios, utilizamos la función **summary**, que muestra el resumen de los datos (en el siguiente comando se muestra un resumen de las columnas 2 y 3):

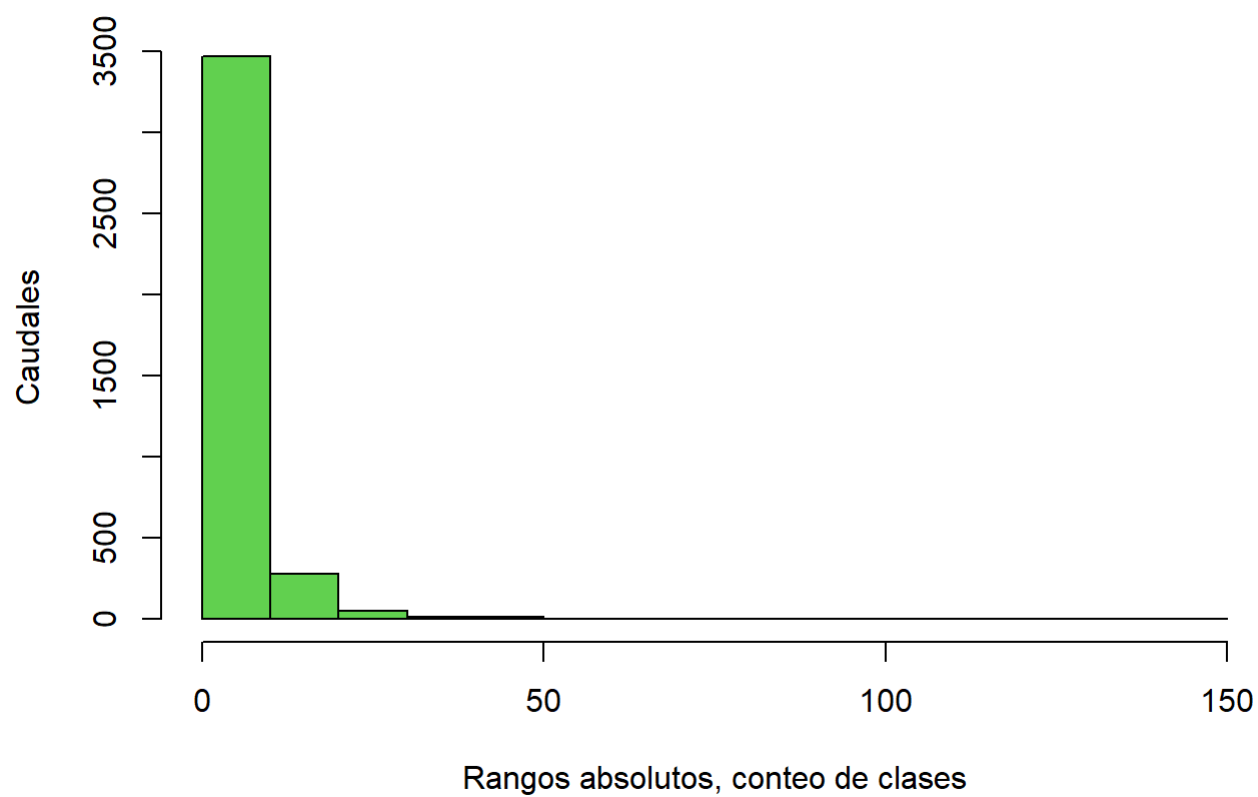
```
summary(inp[,2:3])
```

```
## Pandora.mm..637km2. Banano.mm..90km2.
## Min.   : 0.180    Min.   : 2.43
## 1st Qu.: 2.390    1st Qu.: 6.88
## Median : 3.590    Median : 10.18
## Mean   : 5.459    Mean   : 14.28
## 3rd Qu.: 5.900    3rd Qu.: 15.46
## Max.   :140.650   Max.   :384.00
```

Para crear un histograma y visualizar como se distribuyen los datos de un mínimo a un máximo, utilizamos la función **hist**:

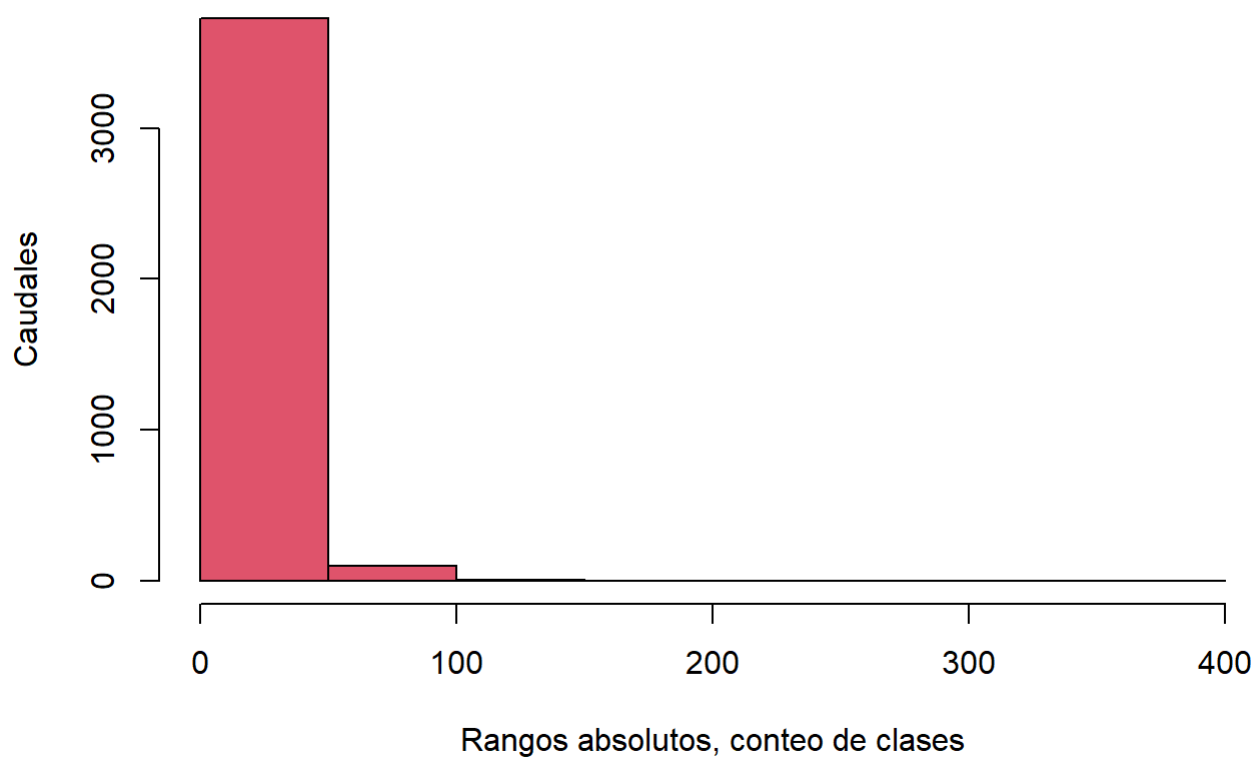
```
hist(inp[,2],col = 3, xlab= "Rangos absolutos, conteo de clases", ylab= "Caudales")
```

Histogram of inp[, 2]



```
hist(inp[,3],col = 2, xlab= "Rangos absolutos, conteo de clases", ylab= "Caudales")
```

Histogram of inp[, 3]

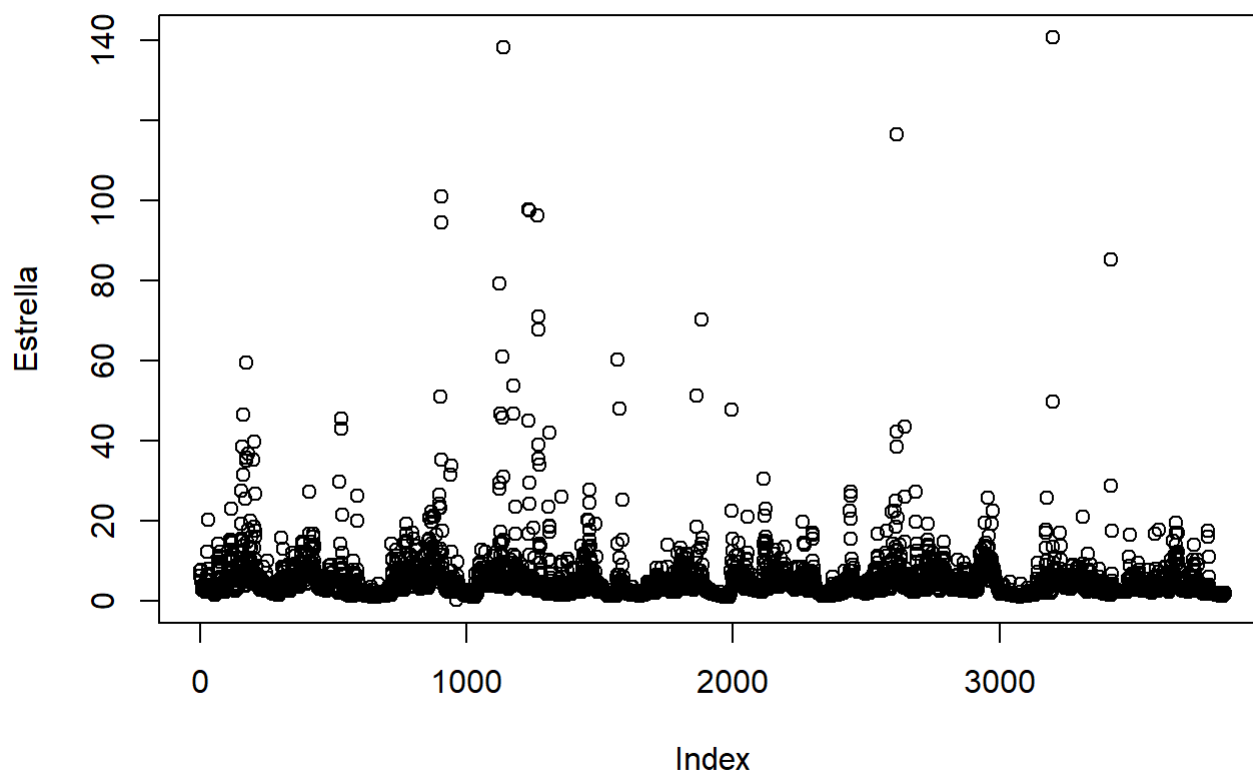


Para no trabajar con columnas o el “inp”, utilizamos la funcion *names* y de esta forma se pueden renombrar: (*nota: la letra **c** significa concatenar y el **attach** es para utilizar directamente los nombres, se debe utilizar con cuidado).

```
names(inp) <- c("fecha", "Estrella", "Banano")  
attach(inp)
```

Plot de ejemplo para verificar renombre de (inp,2):

```
plot(Estrella)
```



Series de tiempo de efectos de balances hidrologicos de años y meses:

Primero se debe definir el formato de la fecha y el tiempo. Se crea un archivo intermedio y creamos la funcion **strptime**:

```
Tempdate <- strptime(inp[,1], format= "%d/%m/%Y")
```

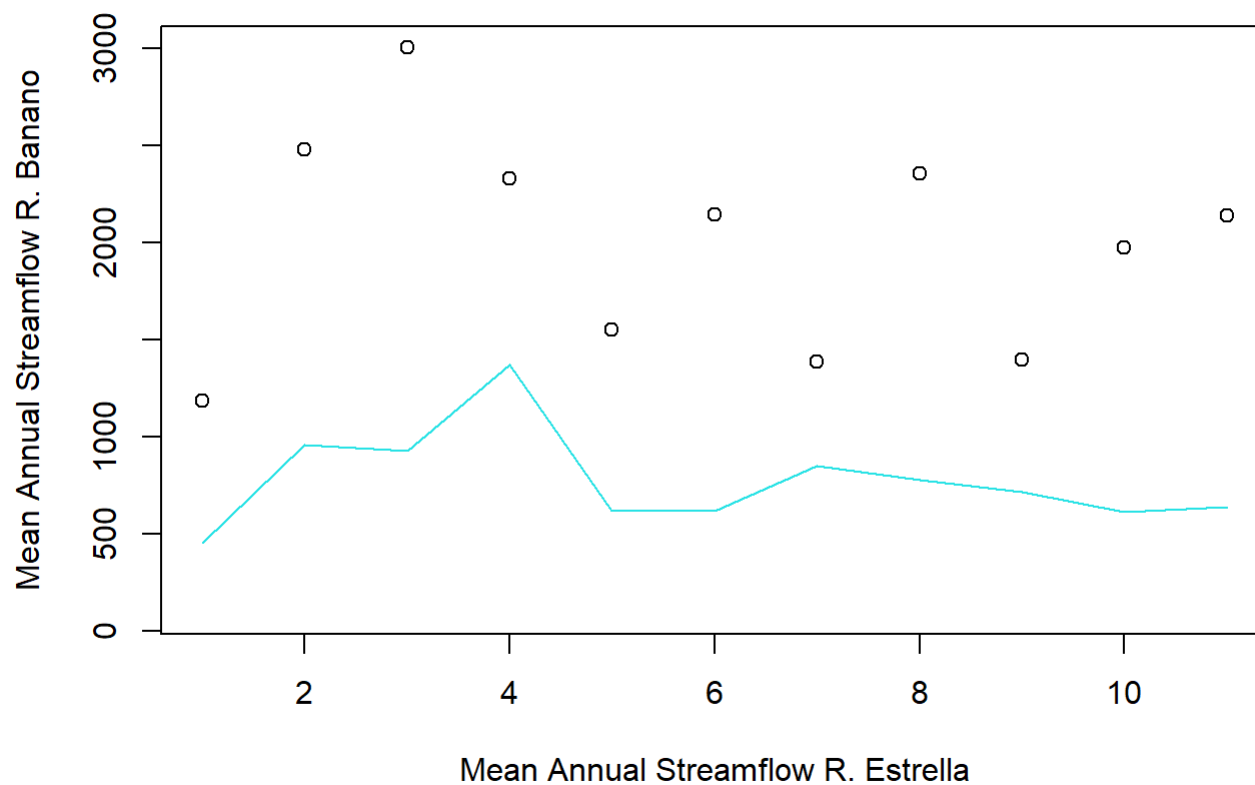
La funcion vectorizada **tapply**, se utiliza para trabajar series de tiempo. La utilizamos para agregar la serie del rio Estrella y rio Banano a promedios anuales; El **write.csv**, se utiliza para crear un archivo con los datos:

Mean Annual Streamflow R. Estrella y R.Banano

```
MAQ_Estrella <- tapply(Estrella, format(Tempdate, format ="%Y"), FUN =sum)
MAQ_Banano <- tapply(Banano, format(Tempdate, format ="%Y"), FUN =sum)
write.csv(rbind(MAQ_Estrella,MAQ_Banano), file ="MAQ.csv")
```

Grafico de valores anuales de caudal: Para ajustar los limites y de esa forma mostrar la informacion de los dos rios, utilizamos **ylim**.

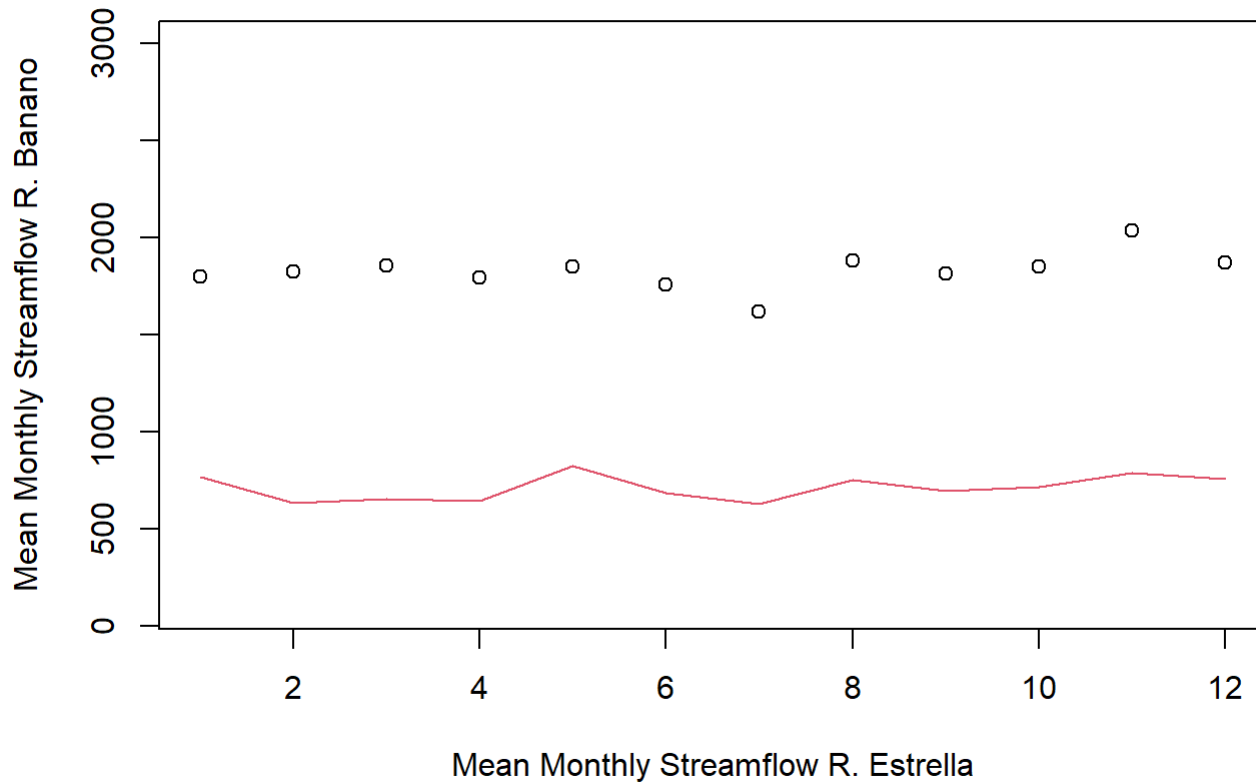
```
plot(MAQ_Banano, ylim=c(100,3000), xlab= "Mean Annual Streamflow R. Estrella", ylab= "Mean Annual Streamflow R. Banano")
lines(MAQ_Estrella, col=5)
```



Aplicamos otravez la funcion vectorizada **tapply**, para agregar la serie del rio Estrella y rio Banano a promedios mensuales:

```
MMQ_Estrella <- tapply(Estrella, format(Tempdate, format = "%m"), FUN = sum)
MMQ_Banano <- tapply(Banano, format(Tempdate, format = "%m"), FUN = sum)
```

```
plot(MMQ_Banano, ylim=c(100,3000), xlab= "Mean Monthly Streamflow R. Estrella", ylab= "Mean Monthly Streamflow R. Banano")
lines(MMQ_Estrella, col=2)
```



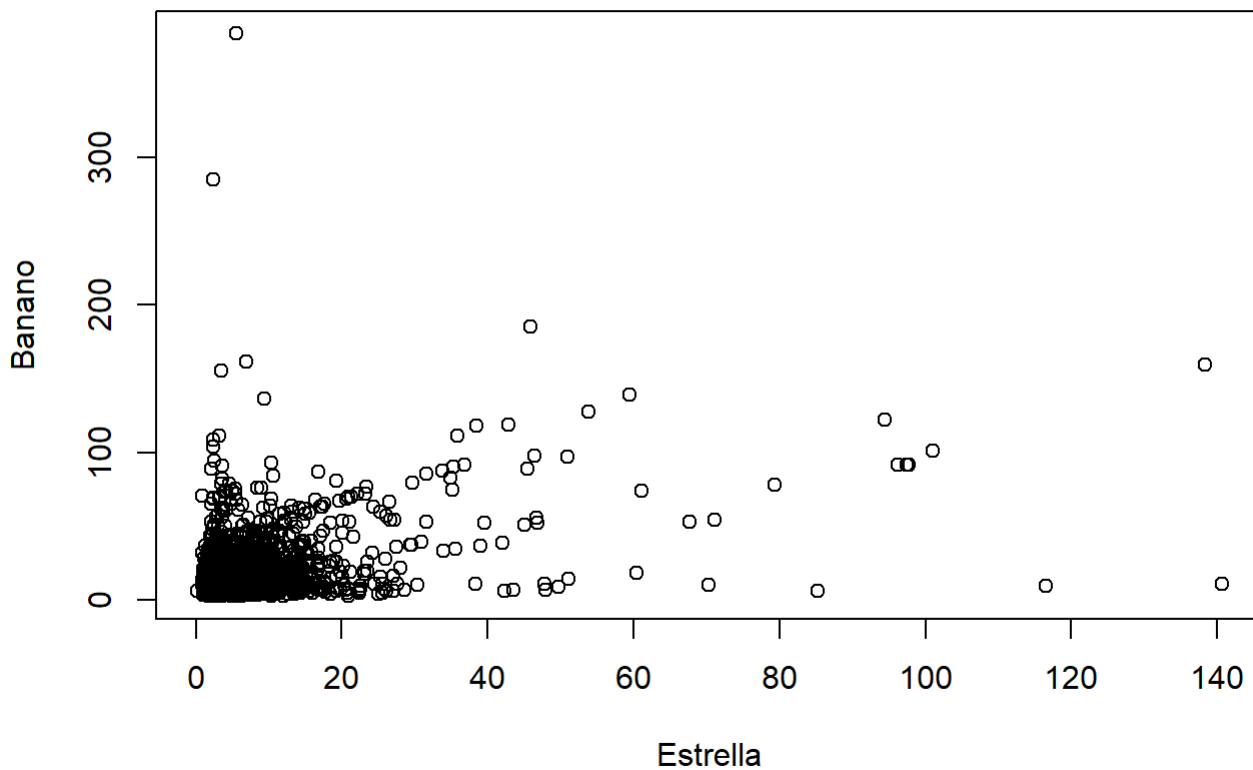
Analisis de correlacion

Creamos un analisis de correlacion para calcular una regresion de datos explorativa usando la funcion **cor** y utilizamos el metodo **spearman** que es un metodo no parametrico para datos ambientales:

```
corinp <- cor(inp[,2:3], method= "spearman")
```

Con **plot** visualizamos el grafico de correlacion:

```
plot(Estrella, Banano)
```



Modelo de regresion lineal

Para calcular un modelo de regresion lineal (cuantificacion matematica de la correlacion), se utiliza la funcion **lm**:

```
inp.lm <-lm(inp[,2] ~ inp[,3], data=inp)
summary(inp.lm)
```

```
##
## Call:
## lm(formula = inp[, 2] ~ inp[, 3], data = inp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -71.297  -2.267  -1.155   0.774  135.889
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.70503    0.15174   17.83  <2e-16 ***
## inp[, 3]      0.19290    0.00707   27.28  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.025 on 3843 degrees of freedom
## Multiple R-squared:  0.1623, Adjusted R-squared:  0.162
## F-statistic: 744.4 on 1 and 3843 DF,  p-value: < 2.2e-16
```


Para visualizar los datos de la regresión lineal en gráficos utilizamos **plot**:

```
plot(inp.lm)
```

