

# INFORME ANALITICA DE DATOS TPAD-1

TPAD-1

12/2/2021

```
#INSTALAMOS LAS LIBRERIAS A UTILIZAR
```

```
install.packages("hexbin")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
install.packages("readr")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
install.packages("ggplot2")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
install.packages("dplyr")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
install.packages("scatterplot3d",dependencies=T)
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
install.packages("tidyverse")
```

```
## Installing package into '/home/rstudio-user/R/x86_64-pc-linux-gnu-library/4.0'  
## (as 'lib' is unspecified)
```

```
#CARGAMOS LAS LIBRERIAS QUE VAMOS A UTILIZAR
```

```
library(readr)
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##     filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##     intersect, setdiff, setequal, union
```

```

library(ggplot2)
library(hexbin)
library(scatterplot3d)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v tibble 3.0.6      v stringr 1.4.0
## v tidyr 1.1.2      v forcats 0.5.1
## v purrr 0.3.4

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

#OBTENEMOS EL ENLACE DE LA CARPETA EN LA QUE ESTAMOS TRABAJANDO
getwd()

## [1] "/cloud/project"

#BUSCAMOS LA RUTA DEL DATAFRAME
cafe_df<-read.csv("/cloud/project/R-MARKDOWN/cafesito/produccion.csv", header = TRUE, sep=",", dec=".")

#NOS INDICA EL TIPO O CLASE DE LA VARIABLE RENDIMIENTO
class(cafe_df$Rendimiento..ha.ton.)

## [1] "numeric"

#A CONTINUACION PODEMOS IDENTIFICAR LA ESTRUCTURA DEL DATAFRAME
str(cafe_df)

## 'data.frame': 266 obs. of 8 variables:
## $ Anio : int 2007 2007 2007 2007 2007 2007 2007 2007 2007 ...
## $ Departamento : chr "ANTIOQUIA" "BOLIVAR" "BOYACA" "CALDAS" ...
## $ Producto : chr "CAFE" "CAFE" "CAFE" "CAFE" ...
## $ Area..ha. : chr "112,343.60" "502" "11,374.50" "78,393.65" ...
## $ Produccion..ton. : chr "120,500.80" "446" "9,683.10" "92,815.00" ...
## $ Rendimiento..ha.ton. : num 1.07 0.89 0.85 1.18 0.93 0.79 0.96 0.57 0.71 0.78 ...
## $ Produccion.Nacional..ton.: num 14.54 0.05 1.17 11.2 0.26 ...
## $ Area.Nacional..ha. : num 14.66 0.07 1.48 10.23 0.3 ...

#CONOCER LA DIMENSION DEL DATAFRAME
dim(cafe_df)

## [1] 266 8

#PODEMOS VER LA INFORMACION RESUMIDA DEL DATAFRAME DE LAS VARIABLES CUANTI-
TATIVAS Y VALORES ESTADISTICOS
summary(cafe_df)

## Anio Departamento Producto Area..ha.
## Min. :2007 Length:266 Length:266 Length:266
## 1st Qu.:2010 Class :character Class :character Class :character
## Median :2012 Mode :character Mode :character Mode :character
## Mean :2012
## 3rd Qu.:2015

```

```
## Max. :2018
## Produccion..ton. Rendimiento..ha.ton. Produccion.Nacional..ton.
## Length:266 Min. :0.0000 Min. : 0.0000
## Class :character 1st Qu.:0.7500 1st Qu.: 0.3525
## Mode :character Median :0.9400 Median : 2.7200
## Mean :0.9364 Mean : 4.5113
## 3rd Qu.:1.1200 3rd Qu.: 7.1475
## Max. :2.0000 Max. :18.6700
## Area.Nacional..ha.
## Min. : 0.000
## 1st Qu.: 0.390
## Median : 3.120
## Mean : 4.511
## 3rd Qu.: 6.875
## Max. :16.430
```

# #INFORMACION CUANTITATIVA DEL DATAFRAME USANDO VARIAS VARIABLES

```
tapply(cafe_df$Rendimiento..ha.ton., cafe_df$Anio, summary)
```

```
## $`2007`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.4500 0.7900 0.9000 0.9505 1.1525 1.5300
##
## $`2008`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.4500 0.7775 0.9050 0.9827 1.2000 1.7900
##
## $`2009`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.3000 0.7600 0.9300 0.8814 1.1125 1.2100
##
## $`2010`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.0000 0.7050 0.9600 0.9061 1.1250 1.5200
##
## $`2011`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.4700 0.6100 0.9000 0.8543 1.0550 1.2000
##
## $`2012`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.0000 0.7450 0.8300 0.8587 0.9150 2.0000
##
## $`2013`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.6000 0.6000 0.7550 0.7595 0.8800 0.9900
##
## $`2014`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.6400 0.6500 0.8150 0.8223 0.9500 1.0600
##
## $`2015`
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.770 0.935 1.065 1.025 1.107 1.150
```

```
##
## $`2016`
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    0.790   0.960   1.120   1.064   1.150   1.190
##
## $`2017`
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    0.660   0.845   1.090   1.068   1.290   1.500
##
## $`2018`
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    0.6200  0.8575  1.1200  1.0795  1.3125  1.5200

#ALGUNOS VALORES ESTADISTICOS
sum(caffe_df$Rendimiento..ha.ton.)

## [1] 249.09
prod(caffe_df$Area.Nacional..ha.)

## [1] 0
var(caffe_df$Rendimiento..ha.ton.)

## [1] 0.07135814
length(caffe_df$Area.Nacional..ha.)

## [1] 266
mean(caffe_df$Rendimiento..ha.ton.)

## [1] 0.9364286
max(caffe_df$Rendimiento..ha.ton.)

## [1] 2
min(caffe_df$Rendimiento..ha.ton.)

## [1] 0
sd(caffe_df$Rendimiento..ha.ton.)

## [1] 0.2671294

#RESTAR EL VALOR MAXIMO Y EL VALOR MINIMO DE LA VARIABLE RENDIMIENTO
max(caffe_df$Rendimiento..ha.ton.)-min(caffe_df$Rendimiento..ha.ton.)

## [1] 2

#NOMBRE DE LAS VARIABLES DEL DATAFRAME
names(caffe_df)

## [1] "Anio"
## [3] "Producto"
## [5] "Produccion..ton."
## [7] "Produccion.Nacional..ton."

"Departamento"
"Area..ha."
"Rendimiento..ha.ton."
"Area.Nacional..ha."

colnames(caffe_df)
```

```
## [1] "Anio" "Departamento"
## [3] "Producto" "Area..ha."
## [5] "Produccion..ton." "Rendimiento..ha.ton."
## [7] "Produccion.Nacional..ton." "Area.Nacional..ha."
```

#VER LAS PRIMERAS Y ULTIMAS 6 FILAS DEL DATAFRAME

```
head(caffe_df)
```

```
##   Anio Departamento Producto Area..ha. Produccion..ton. Rendimiento..ha.ton.
## 1 2007     ANTIOQUIA   CAFE 112,343.60      120,500.80          1.07
## 2 2007     BOLIVAR    CAFE      502          446          0.89
## 3 2007     BOYACA     CAFE 11,374.50      9,683.10          0.85
## 4 2007     CALDAS     CAFE 78,393.65     92,815.00          1.18
## 5 2007     CAQUETA    CAFE  2,295.00     2,134.00          0.93
## 6 2007     CASANARE   CAFE  2,605.00     2,048.40          0.79
##   Produccion.Nacional..ton. Area.Nacional..ha.
## 1              14.54              14.66
## 2              0.05              0.07
## 3              1.17              1.48
## 4             11.20             10.23
## 5              0.26              0.30
## 6              0.25              0.34
```

```
tail(caffe_df)
```

```
##   Anio Departamento Producto Area..ha. Produccion..ton.
## 261 2018     PUTUMAYO   CAFE   209.93      289.5
## 262 2018     QUINDIO    CAFE 16,374.73     17,739.03
## 263 2018     RISARALDA   CAFE 35,874.73     45,918.75
## 264 2018     SANTANDER   CAFE 42,269.07     55,918.71
## 265 2018     TOLIMA     CAFE 97,304.04     97,451.31
## 266 2018 VALLE DEL CAUCA CAFE 48,305.31     49,667.88
##   Rendimiento..ha.ton. Produccion.Nacional..ton. Area.Nacional..ha.
## 261              1.38              0.03              0.03
## 262              1.08              2.07              2.21
## 263              1.28              5.37              4.83
## 264              1.32              6.53              5.69
## 265              1.00             11.39             13.11
## 266              1.03              5.80              6.51
```

#INSPECCIONAR DE FORMA DETALLADA EN EL DATAFRAME SELECCIONANDO SUS COLUMNAS Y FILAS

```
caffe_df[3]
```

```
##   Producto
## 1      CAFE
## 2      CAFE
## 3      CAFE
## 4      CAFE
## 5      CAFE
## 6      CAFE
## 7      CAFE
## 8      CAFE
## 9      CAFE
## 10     CAFE
```

##	11	CAFE
##	12	CAFE
##	13	CAFE
##	14	CAFE
##	15	CAFE
##	16	CAFE
##	17	CAFE
##	18	CAFE
##	19	CAFE
##	20	CAFE
##	21	CAFE
##	22	CAFE
##	23	CAFE
##	24	CAFE
##	25	CAFE
##	26	CAFE
##	27	CAFE
##	28	CAFE
##	29	CAFE
##	30	CAFE
##	31	CAFE
##	32	CAFE
##	33	CAFE
##	34	CAFE
##	35	CAFE
##	36	CAFE
##	37	CAFE
##	38	CAFE
##	39	CAFE
##	40	CAFE
##	41	CAFE
##	42	CAFE
##	43	CAFE
##	44	CAFE
##	45	CAFE
##	46	CAFE
##	47	CAFE
##	48	CAFE
##	49	CAFE
##	50	CAFE
##	51	CAFE
##	52	CAFE
##	53	CAFE
##	54	CAFE
##	55	CAFE
##	56	CAFE
##	57	CAFE
##	58	CAFE
##	59	CAFE
##	60	CAFE
##	61	CAFE
##	62	CAFE
##	63	CAFE
##	64	CAFE

## 65	CAFE
## 66	CAFE
## 67	CAFE
## 68	CAFE
## 69	CAFE
## 70	CAFE
## 71	CAFE
## 72	CAFE
## 73	CAFE
## 74	CAFE
## 75	CAFE
## 76	CAFE
## 77	CAFE
## 78	CAFE
## 79	CAFE
## 80	CAFE
## 81	CAFE
## 82	CAFE
## 83	CAFE
## 84	CAFE
## 85	CAFE
## 86	CAFE
## 87	CAFE
## 88	CAFE
## 89	CAFE
## 90	CAFE
## 91	CAFE
## 92	CAFE
## 93	CAFE
## 94	CAFE
## 95	CAFE
## 96	CAFE
## 97	CAFE
## 98	CAFE
## 99	CAFE
## 100	CAFE
## 101	CAFE
## 102	CAFE
## 103	CAFE
## 104	CAFE
## 105	CAFE
## 106	CAFE
## 107	CAFE
## 108	CAFE
## 109	CAFE
## 110	CAFE
## 111	CAFE
## 112	CAFE
## 113	CAFE
## 114	CAFE
## 115	CAFE
## 116	CAFE
## 117	CAFE
## 118	CAFE

## 119	CAFE
## 120	CAFE
## 121	CAFE
## 122	CAFE
## 123	CAFE
## 124	CAFE
## 125	CAFE
## 126	CAFE
## 127	CAFE
## 128	CAFE
## 129	CAFE
## 130	CAFE
## 131	CAFE
## 132	CAFE
## 133	CAFE
## 134	CAFE
## 135	CAFE
## 136	CAFE
## 137	CAFE
## 138	CAFE
## 139	CAFE
## 140	CAFE
## 141	CAFE
## 142	CAFE
## 143	CAFE
## 144	CAFE
## 145	CAFE
## 146	CAFE
## 147	CAFE
## 148	CAFE
## 149	CAFE
## 150	CAFE
## 151	CAFE
## 152	CAFE
## 153	CAFE
## 154	CAFE
## 155	CAFE
## 156	CAFE
## 157	CAFE
## 158	CAFE
## 159	CAFE
## 160	CAFE
## 161	CAFE
## 162	CAFE
## 163	CAFE
## 164	CAFE
## 165	CAFE
## 166	CAFE
## 167	CAFE
## 168	CAFE
## 169	CAFE
## 170	CAFE
## 171	CAFE
## 172	CAFE



## 173	CAFE
## 174	CAFE
## 175	CAFE
## 176	CAFE
## 177	CAFE
## 178	CAFE
## 179	CAFE
## 180	CAFE
## 181	CAFE
## 182	CAFE
## 183	CAFE
## 184	CAFE
## 185	CAFE
## 186	CAFE
## 187	CAFE
## 188	CAFE
## 189	CAFE
## 190	CAFE
## 191	CAFE
## 192	CAFE
## 193	CAFE
## 194	CAFE
## 195	CAFE
## 196	CAFE
## 197	CAFE
## 198	CAFE
## 199	CAFE
## 200	CAFE
## 201	CAFE
## 202	CAFE
## 203	CAFE
## 204	CAFE
## 205	CAFE
## 206	CAFE
## 207	CAFE
## 208	CAFE
## 209	CAFE
## 210	CAFE
## 211	CAFE
## 212	CAFE
## 213	CAFE
## 214	CAFE
## 215	CAFE
## 216	CAFE
## 217	CAFE
## 218	CAFE
## 219	CAFE
## 220	CAFE
## 221	CAFE
## 222	CAFE
## 223	CAFE
## 224	CAFE
## 225	CAFE
## 226	CAFE

```
## 227    CAFE
## 228    CAFE
## 229    CAFE
## 230    CAFE
## 231    CAFE
## 232    CAFE
## 233    CAFE
## 234    CAFE
## 235    CAFE
## 236    CAFE
## 237    CAFE
## 238    CAFE
## 239    CAFE
## 240    CAFE
## 241    CAFE
## 242    CAFE
## 243    CAFE
## 244    CAFE
## 245    CAFE
## 246    CAFE
## 247    CAFE
## 248    CAFE
## 249    CAFE
## 250    CAFE
## 251    CAFE
## 252    CAFE
## 253    CAFE
## 254    CAFE
## 255    CAFE
## 256    CAFE
## 257    CAFE
## 258    CAFE
## 259    CAFE
## 260    CAFE
## 261    CAFE
## 262    CAFE
## 263    CAFE
## 264    CAFE
## 265    CAFE
## 266    CAFE
```

```
cafe_df[,4]
```

```
##   [1] "112,343.60" "502"      "11,374.50"  "78,393.65"  "2,295.00"
##   [6] "2,605.00"   "53,471.00"  "23,172.00"  "290"        "43,017.30"
##  [11] "89,661.56"   "4,785.00"   "17,506.00"  "2,048.00"   "24,458.50"
##  [16] "30,171.84"   "35"         "19,904.00"  "47,689.25"  "34,406.67"
##  [21] "91,679.10"   "76,667.80"  "114,694.00" "572"        "10,778.50"
##  [26] "74,897.00"   "2,735.00"   "2,149.00"   "56,208.00"  "23,198.00"
##  [31] "90"          "43,633.35"  "89,131.20"  "4,553.00"   "17,521.00"
##  [36] "2,146.00"    "25,582.00"  "30,171.84"  "31"         "19,571.00"
##  [41] "47,227.00"   "34,169.37"  "86,829.20"  "72,419.00"  "112,420.20"
##  [46] "770"         "10,672.50"  "73,083.00"  "2,332.00"   "1,904.00"
##  [51] "57,860.00"   "23,420.00"  "70"         "43,475.84"  "86,726.78"
##  [56] "4,488.00"    "17,036.00"  "2,216.00"   "26,467.20"  "33,552.58"
```

```

## [61] "23"          "19,052.00"  "45,428.00"  "37,985.90"  "88,667.00"
## [66] "67,001.30"  "111,602.71" "0"          "850"        "9,427.00"
## [71] "72,240.58"  "2,536.00"   "2,198.00"   "55,162.00"  "22,489.50"
## [76] "157.5"      "44,264.16"  "87,139.53"  "4,207.00"   "17,000.00"
## [81] "2,326.00"   "23,504.05"  "30,731.96"  "24"         "18,159.00"
## [86] "47,308.00"  "39,000.64"  "84,658.70"  "69,332.10"  "106,419.57"
## [91] "10"         "850"        "8,441.74"   "66,331.61"  "2,810.00"
## [96] "2,081.50"   "54,246.42"  "22,350.00"  "157.5"      "37,478.87"
## [101] "78,792.21"  "4,100.00"   "16,577.00"  "2,578.00"   "24,263.80"
## [106] "21,520.45"  "40"         "20,139.30"  "44,733.64"  "37,282.04"
## [111] "93,145.35"  "68,038.40"  "112,221.14" "870"        "6,698.20"
## [116] "54,871.88"  "2,882.50"   "2,322.00"   "56,825.00"  "22,911.00"
## [121] "70"         "37,175.06"  "0"          "79,809.34"  "5,143.00"
## [126] "17,686.00"  "2,783.00"   "27,806.40"  "19,339.31"  "42"
## [131] "21,109.83"  "45,588.03"  "33,947.15"  "90,904.48"  "69,456.71"
## [136] "109,755.50" "659.04"     "9,289.05"   "60,264.29"  "2,905.84"
## [141] "2,232.94"   "74,105.64"  "25,106.39"  "125.01"     "36,189.18"
## [146] "118,200.88" "5,750.70"   "17,016.72"  "2,483.43"   "32,136.51"
## [151] "25,332.45"  "24.27"      "21,203.03"  "39,615.60"  "38,613.68"
## [156] "97,308.81"  "53,481.02"  "110,115.86" "936.34"     "9,834.39"
## [161] "59,757.18"  "3,074.92"   "2,599.43"   "77,068.46"  "26,138.58"
## [166] "136.88"     "33,623.54"  "128,273.15" "6,078.64"   "18,533.11"
## [171] "2,739.71"   "33,608.32"  "23,724.20"  "101.16"     "21,462.81"
## [176] "40,154.46"  "40,733.20"  "100,832.91" "56,035.94"  "109,649.61"
## [181] "1,065.07"   "10,461.85"  "58,376.40"  "3,410.56"   "2,752.31"
## [186] "77,405.83"  "25,948.50"  "137.47"     "34,101.49"  "130,452.40"
## [191] "5,631.53"   "17,996.31"  "2,922.21"   "33,490.93"  "22,940.64"
## [196] "128.65"     "21,491.21"  "41,732.03"  "42,679.11"  "103,368.73"
## [201] "54,938.79"  "105,666.60" "1,065.97"   "10,181.80"  "56,022.04"
## [206] "3,392.22"   "2,671.04"   "78,421.95"  "25,530.59"  "134.96"
## [211] "33,214.17"  "126,052.15" "5,531.20"   "17,745.80"  "2,924.89"
## [216] "32,750.16"  "21,520.64"  "20,041.70"  "40,472.26"  "41,387.79"
## [221] "100,328.77" "52,648.25"  "99,311.53"  "1,137.42"   "9,598.33"
## [226] "51,854.59"  "3,408.69"   "2,436.63"   "80,289.56"  "25,158.80"
## [231] "125.67"     "30,894.16"  "122,575.76" "5,340.80"   "18,129.50"
## [236] "2,926.85"   "33,639.55"  "21,409.77"  "209.29"     "17,699.67"
## [241] "37,334.16"  "42,327.26"  "96,018.89"  "51,470.86"  "98,038.15"
## [246] "1,182.13"   "9,653.45"   "50,762.22"  "3,485.24"   "2,360.55"
## [251] "82,085.54"  "23,915.45"  "140.33"     "29,085.24"  "122,002.46"
## [256] "4,810.97"   "17,414.32"  "2,761.01"   "33,465.54"  "20,873.04"
## [261] "209.93"     "16,374.73"  "35,874.73"  "42,269.07"  "97,304.04"
## [266] "48,305.31"

```

```
cafe_df[,2]
```

```

## [1] "ANTIOQUIA"      "BOLIVAR"      "BOYACA"
## [4] "CALDAS"         "CAQUETA"      "CASANARE"
## [7] "CAUCA"          "CESAR"        "CHOCO"
## [10] "CUNDINAMARCA"  "HUILA"        "LA GUAJIRA"
## [13] "MAGDALENA"     "META"         "NARIÑO"
## [16] "NORTE DE SANTANDER" "PUTUMAYO"     "QUINDIO"
## [19] "RISARALDA"     "SANTANDER"    "TOLIMA"
## [22] "VALLE DEL CAUCA" "ANTIOQUIA"    "BOLIVAR"
## [25] "BOYACA"        "CALDAS"       "CAQUETA"
## [28] "CASANARE"      "CAUCA"        "CESAR"

```

## [31]	"CHOCO"	"CUNDINAMARCA"	"HUILA"
## [34]	"LA GUAJIRA"	"MAGDALENA"	"META"
## [37]	"NARIÑO"	"NORTE DE SANTANDER"	"PUTUMAYO"
## [40]	"QUINDIO"	"RISARALDA"	"SANTANDER"
## [43]	"TOLIMA"	"VALLE DEL CAUCA"	"ANTIOQUIA"
## [46]	"BOLIVAR"	"BOYACA"	"CALDAS"
## [49]	"CAQUETA"	"CASANARE"	"CAUCA"
## [52]	"CESAR"	"CHOCO"	"CUNDINAMARCA"
## [55]	"HUILA"	"LA GUAJIRA"	"MAGDALENA"
## [58]	"META"	"NARIÑO"	"NORTE DE SANTANDER"
## [61]	"PUTUMAYO"	"QUINDIO"	"RISARALDA"
## [64]	"SANTANDER"	"TOLIMA"	"VALLE DEL CAUCA"
## [67]	"ANTIOQUIA"	"ARAUCA"	"BOLIVAR"
## [70]	"BOYACA"	"CALDAS"	"CAQUETA"
## [73]	"CASANARE"	"CAUCA"	"CESAR"
## [76]	"CHOCO"	"CUNDINAMARCA"	"HUILA"
## [79]	"LA GUAJIRA"	"MAGDALENA"	"META"
## [82]	"NARIÑO"	"NORTE DE SANTANDER"	"PUTUMAYO"
## [85]	"QUINDIO"	"RISARALDA"	"SANTANDER"
## [88]	"TOLIMA"	"VALLE DEL CAUCA"	"ANTIOQUIA"
## [91]	"ARAUCA"	"BOLIVAR"	"BOYACA"
## [94]	"CALDAS"	"CAQUETA"	"CASANARE"
## [97]	"CAUCA"	"CESAR"	"CHOCO"
## [100]	"CUNDINAMARCA"	"HUILA"	"LA GUAJIRA"
## [103]	"MAGDALENA"	"META"	"NARIÑO"
## [106]	"NORTE DE SANTANDER"	"PUTUMAYO"	"QUINDIO"
## [109]	"RISARALDA"	"SANTANDER"	"TOLIMA"
## [112]	"VALLE DEL CAUCA"	"ANTIOQUIA"	"BOLIVAR"
## [115]	"BOYACA"	"CALDAS"	"CAQUETA"
## [118]	"CASANARE"	"CAUCA"	"CESAR"
## [121]	"CHOCO"	"CUNDINAMARCA"	"GUAVIARE"
## [124]	"HUILA"	"LA GUAJIRA"	"MAGDALENA"
## [127]	"META"	"NARIÑO"	"NORTE DE SANTANDER"
## [130]	"PUTUMAYO"	"QUINDIO"	"RISARALDA"
## [133]	"SANTANDER"	"TOLIMA"	"VALLE DEL CAUCA"
## [136]	"ANTIOQUIA"	"BOLIVAR"	"BOYACA"
## [139]	"CALDAS"	"CAQUETA"	"CASANARE"
## [142]	"CAUCA"	"CESAR"	"CHOCO"
## [145]	"CUNDINAMARCA"	"HUILA"	"LA GUAJIRA"
## [148]	"MAGDALENA"	"META"	"NARIÑO"
## [151]	"NORTE DE SANTANDER"	"PUTUMAYO"	"QUINDIO"
## [154]	"RISARALDA"	"SANTANDER"	"TOLIMA"
## [157]	"VALLE DEL CAUCA"	"ANTIOQUIA"	"BOLIVAR"
## [160]	"BOYACA"	"CALDAS"	"CAQUETA"
## [163]	"CASANARE"	"CAUCA"	"CESAR"
## [166]	"CHOCO"	"CUNDINAMARCA"	"HUILA"
## [169]	"LA GUAJIRA"	"MAGDALENA"	"META"
## [172]	"NARIÑO"	"NORTE DE SANTANDER"	"PUTUMAYO"
## [175]	"QUINDIO"	"RISARALDA"	"SANTANDER"
## [178]	"TOLIMA"	"VALLE DEL CAUCA"	"ANTIOQUIA"
## [181]	"BOLIVAR"	"BOYACA"	"CALDAS"
## [184]	"CAQUETA"	"CASANARE"	"CAUCA"
## [187]	"CESAR"	"CHOCO"	"CUNDINAMARCA"
## [190]	"HUILA"	"LA GUAJIRA"	"MAGDALENA"

```
## [193] "META"           "NARIÑO"           "NORTE DE SANTANDER"
## [196] "PUTUMAYO"        "QUINDIO"          "RISARALDA"
## [199] "SANTANDER"       "TOLIMA"           "VALLE DEL CAUCA"
## [202] "ANTIOQUIA"       "BOLIVAR"          "BOYACA"
## [205] "CALDAS"          "CAQUETA"          "CASANARE"
## [208] "CAUCA"           "CESAR"            "CHOCO"
## [211] "CUNDINAMARCA"    "HUILA"            "LA GUAJIRA"
## [214] "MAGDALENA"       "META"             "NARIÑO"
## [217] "NORTE DE SANTANDER" "QUINDIO"          "RISARALDA"
## [220] "SANTANDER"       "TOLIMA"           "VALLE DEL CAUCA"
## [223] "ANTIOQUIA"       "BOLIVAR"          "BOYACA"
## [226] "CALDAS"          "CAQUETA"          "CASANARE"
## [229] "CAUCA"           "CESAR"            "CHOCO"
## [232] "CUNDINAMARCA"    "HUILA"            "LA GUAJIRA"
## [235] "MAGDALENA"       "META"             "NARIÑO"
## [238] "NORTE DE SANTANDER" "PUTUMAYO"         "QUINDIO"
## [241] "RISARALDA"       "SANTANDER"        "TOLIMA"
## [244] "VALLE DEL CAUCA" "ANTIOQUIA"        "BOLIVAR"
## [247] "BOYACA"           "CALDAS"           "CAQUETA"
## [250] "CASANARE"        "CAUCA"            "CESAR"
## [253] "CHOCO"           "CUNDINAMARCA"     "HUILA"
## [256] "LA GUAJIRA"      "MAGDALENA"        "META"
## [259] "NARIÑO"          "NORTE DE SANTANDER" "PUTUMAYO"
## [262] "QUINDIO"         "RISARALDA"        "SANTANDER"
## [265] "TOLIMA"          "VALLE DEL CAUCA"
```

```
cafe_df[2,]
```

```
## Anio Departamento Producto Area..ha. Produccion..ton. Rendimiento..ha.ton.
## 2 2007 BOLIVAR CAFE 502 446 0.89
## Produccion.Nacional..ton. Area.Nacional..ha.
## 2 0.05 0.07
```

```
cafe_df[100,]
```

```
## Anio Departamento Producto Area..ha. Produccion..ton. Rendimiento..ha.ton.
## 100 2011 CUNDINAMARCA CAFE 37,478.87 32,780.35 0.87
## Produccion.Nacional..ton. Area.Nacional..ha.
## 100 5.12 5.26
```

#CON LA FUNCION TABLE LISTAMOS LAS FILAS DE LA VARIABLE SELECCIONADA

```
table(cafe_df$Produccion..ton.)
```

```
##
##      0  1,089.74  1,128.32  1,338.56  1,388.13  1,617.20  1,629.25
##      2      1      1      1      1      1      1
## 1,650.41  1,656.96  1,672.60  1,688.60  1,718.25  1,747.51  1,933.00
##      1      1      1      1      1      1      1
## 1,950.84 10,200.84 10,221.69 10,826.24 101,201.88 102,147.00 102,403.24
##      1      1      1      1      1      1      1
## 103,703.00 104,336.56 104,609.42 105,563.88 105,976.19 105.93 11,035.85
##      1      1      1      1      1      1      1
## 11,937.90 111,452.91 112,322.38 113,505.20 115,267.98 115,874.98 119,970.68
##      1      1      1      1      1      1      1
##      12 12,012.98 12,214.54 12,332.00 12,770.00 120,365.77 120,500.80
##      1      1      1      1      1      1      1
```

##	121,253.38	124.67	125.42	129,052.51	13,276.08	13,278.50	13,301.60
##	1	1	1	1	1	1	1
##	13,412.80	13,593.24	13,593.25	13,600.00	13,841.45	131,316.47	133,787.95
##	1	1	1	1	1	1	1
##	135,971.20	136,161.86	14,005.00	14,017.00	14,096.05	14,943.62	140
##	1	1	1	1	1	1	1
##	140,398.62	141,898.91	145,154.42	145,168.10	15,050.27	15,108.55	15,185.79
##	1	1	1	1	1	1	1
##	158.2	158.85	16,628.14	16,691.31	16,935.63	16.87	160.62
##	1	1	1	1	1	1	1
##	17,031.09	17,739.03	18,030.13	18,792.05	181.42	19,590.10	19,994.35
##	1	1	1	1	1	1	1
##	2,023.50	2,048.40	2,079.70	2,133.10	2,134.00	2,188.92	2,221.90
##	1	1	1	1	1	1	1
##	2,328.90	2,332.00	2,340.40	2,393.00	2,446.38	2,469.00	2,503.81
##	1	1	1	1	1	1	1
##	2,528.40	2,533.75	2,564.86	2,626.73	2,638.88	2,902.50	2,958.70
##	1	1	1	1	1	1	1
##	2,990.91	20,267.64	20,599.27	20,814.11	205.9	21,065.00	21,985.00
##	1	1	1	1	1	1	1
##	22,089.82	22,111.65	22,240.81	22,518.42	22,649.03	23,271.89	23,409.44
##	1	1	1	1	1	1	1
##	23,471.69	23,669.00	23,791.30	24,073.95	24,594.10	24,694.56	24,993.74
##	1	1	1	1	1	1	1
##	25,118.55	25,426.00	26,311.61	26.7	27,094.16	27,487.71	28,077.94
##	1	1	1	2	1	1	1
##	28,606.96	282.18	289.5	29,016.75	29,469.52	292.6	3,206.35
##	1	1	1	1	1	1	1
##	3,322.42	3,434.30	3,447.31	3,516.80	3,749.27	3,861.63	3,877.62
##	1	1	1	1	1	1	1
##	3,923.80	30,227.02	30,786.41	31,165.15	31,262.50	31,413.34	31,770.05
##	1	1	1	1	1	1	1
##	32,321.56	32,580.24	32,780.35	33,729.14	33,943.39	34	34,512.79
##	1	1	1	1	1	1	1
##	35,004.18	35,679.42	35.6	36,607.56	36,989.43	37,020.90	37,118.07
##	1	1	1	1	1	1	1
##	37,214.80	39,073.92	395.07	4,013.11	4,317.50	4,387.19	4,981.59
##	1	1	1	1	1	1	1
##	41,645.39	42,719.53	42,948.40	446	45,113.00	45,918.75	45.8
##	1	1	1	1	1	1	1
##	46,779.71	47,215.69	47,221.00	47,304.16	47,357.02	47,512.36	48,073.00
##	1	1	1	1	1	1	1
##	48.4	49,042.31	49,667.88	49,799.28	5,108.33	5,280.40	5,591.05
##	1	1	1	1	1	1	1
##	5,643.39	50,588.14	51,348.00	51,687.80	510	53,288.42	53,648.00
##	1	1	1	1	2	1	1
##	54,115.96	54,908.68	55,918.71	56,303.92	57,067.08	57,583.56	58,634.19
##	1	1	1	1	1	1	1
##	6,364.41	60,079.00	606.93	61,190.55	62,711.08	62,869.38	63,365.76
##	1	1	1	1	1	1	1
##	65,475.63	65,666.43	652.5	66,661.14	67,231.37	68	68,668.20
##	1	1	1	1	1	1	1
##	68,670.96	69,496.65	69,618.24	7,083.07	7,638.99	7,780.34	711
##	1	1	1	1	1	1	1

```
## 72,091.00 72,842.55 734.91 748.97 76.04 77,215.36 78,254.77
## 1 1 1 1 1 1 1
## 78,805.87 78.75 8,567.97 81,668.22 83,626.44 85,027.49 85,150.66
## 1 1 1 1 1 1 1
## 85,212.64 86,453.62 86,884.00 87,642.49 88,633.10 9,501.54 9,547.30
## 1 1 1 1 1 1 1
## 9,583.80 9,683.10 91,621.30 92,815.00 94,230.20 94,556.71 95,957.90
## 1 1 1 1 1 1 1
## 97,451.31 97,922.49 98
## 1 1 2
```

```
table(cafe_df$Rendimiento..ha.ton.)
```

```
##
## 0 0.3 0.38 0.45 0.47 0.49 0.51 0.52 0.55 0.57 0.59 0.6 0.62 0.63 0.64 0.65
## 2 1 1 2 1 1 1 1 1 4 2 10 7 1 1 7
## 0.66 0.67 0.69 0.7 0.71 0.72 0.74 0.75 0.76 0.77 0.78 0.79 0.8 0.81 0.82 0.83
## 5 2 5 1 2 2 2 6 2 4 2 5 8 3 4 1
## 0.84 0.85 0.86 0.87 0.88 0.89 0.9 0.91 0.92 0.93 0.94 0.95 0.96 0.97 0.98 0.99
## 1 7 4 2 2 6 2 5 2 3 3 2 5 5 3 5
## 1 1.01 1.02 1.03 1.04 1.05 1.06 1.07 1.08 1.09 1.1 1.11 1.12 1.13 1.14 1.15
## 5 2 2 2 2 4 5 3 5 5 5 4 5 3 4 9
## 1.16 1.17 1.18 1.19 1.2 1.21 1.22 1.23 1.24 1.25 1.26 1.27 1.28 1.29 1.3 1.32
## 3 3 2 4 2 2 2 1 2 1 1 1 2 1 2 2
## 1.33 1.35 1.37 1.38 1.4 1.41 1.44 1.45 1.47 1.5 1.52 1.53 1.79 2
## 1 2 1 1 1 1 1 1 1 1 2 1 1 1
```

```
#PODEMOS VER LA UTILIDAD E INFORMACION DE LA FUNCION
```

```
help (subset)
```

```
#CON LA FUNCION SUBSET PODEMOS SELECCIONAR POR VARIABLES EN NUESTRO
DATAFRAME
```

```
subset(cafe_df, subset = Departamento == 'HUILA')
```

```
## Anio Departamento Producto Area..ha. Produccion..ton. Rendimiento..ha.ton.
## 11 2007 HUILA CAFE 89,661.56 129,052.51 1.44
## 33 2008 HUILA CAFE 89,131.20 131,316.47 1.47
## 55 2009 HUILA CAFE 86,726.78 104,609.42 1.21
## 78 2010 HUILA CAFE 87,139.53 104,336.56 1.20
## 101 2011 HUILA CAFE 78,792.21 85,150.66 1.08
## 124 2012 HUILA CAFE 79,809.34 85,212.64 1.07
## 146 2013 HUILA CAFE 118,200.88 115,874.98 0.98
## 168 2014 HUILA CAFE 128,273.15 135,971.20 1.06
## 190 2015 HUILA CAFE 130,452.40 145,168.10 1.11
## 212 2016 HUILA CAFE 126,052.15 145,154.42 1.15
## 233 2017 HUILA CAFE 122,575.76 133,787.95 1.09
## 255 2018 HUILA CAFE 122,002.46 136,161.86 1.12
## Produccion.Nacional..ton. Area.Nacional..ha.
## 11 15.57 11.70
## 33 15.85 11.75
## 55 14.76 11.49
## 78 13.39 11.71
## 101 13.30 11.06
## 124 13.60 11.23
## 146 17.77 15.31
```

```
## 168          18.67          16.12
## 190          17.07          16.28
## 212          17.00          16.21
## 233          15.71          16.27
## 255          15.91          16.43
```

```
subset(cafe_df, subset = Departamento == 'META')
```

```
##      Anio Departamento Producto Area..ha. Produccion..ton. Rendimiento..ha.ton.
## 14  2007          META    CAFE  2,048.00          1,617.20          0.79
## 36  2008          META    CAFE  2,146.00          1,656.96          0.77
## 58  2009          META    CAFE  2,216.00          1,672.60          0.75
## 81  2010          META    CAFE  2,326.00          2,221.90          0.96
## 104 2011          META    CAFE  2,578.00          2,533.75          0.98
## 127 2012          META    CAFE  2,783.00          2,133.10          0.77
## 149 2013          META    CAFE  2,483.43          1,650.41          0.66
## 171 2014          META    CAFE  2,739.71          1,950.84          0.71
## 193 2015          META    CAFE  2,922.21          3,206.35          1.10
## 215 2016          META    CAFE  2,924.89          3,322.42          1.14
## 236 2017          META    CAFE  2,926.85          4,013.11          1.37
## 258 2018          META    CAFE  2,761.01          3,877.62          1.40
##      Produccion.Nacional..ton. Area.Nacional..ha.
## 14          0.20          0.27
## 36          0.20          0.28
## 58          0.24          0.29
## 81          0.29          0.31
## 104         0.40          0.36
## 127         0.34          0.39
## 149         0.25          0.32
## 171         0.27          0.34
## 193         0.38          0.36
## 215         0.39          0.38
## 236         0.47          0.39
## 258         0.45          0.37
```

```
subset(cafe_df, subset = Anio == '2009')
```

```
##      Anio      Departamento Producto Area..ha. Produccion..ton.
## 45  2009      ANTIOQUIA    CAFE  112,420.20      103,703.00
## 46  2009      BOLIVAR    CAFE      770          292.6
## 47  2009      BOYACA    CAFE  10,672.50          8,567.97
## 48  2009      CALDAS    CAFE  73,083.00          81,668.22
## 49  2009      CAQUETA    CAFE   2,332.00          2,332.00
## 50  2009      CASANARE    CAFE   1,904.00          2,079.70
## 51  2009      CAUCA    CAFE  57,860.00          47,221.00
## 52  2009      CESAR    CAFE  23,420.00          12,770.00
## 53  2009      CHOCO    CAFE      70           78.75
## 54  2009      CUNDINAMARCA CAFE  43,475.84          37,118.07
## 55  2009      HUILA    CAFE  86,726.78          104,609.42
## 56  2009      LA GUAJIRA CAFE   4,488.00          2,340.40
## 57  2009      MAGDALENA CAFE  17,036.00          13,412.80
## 58  2009      META    CAFE   2,216.00          1,672.60
## 59  2009      NARIÑO    CAFE  26,467.20          27,487.71
## 60  2009  NORTE DE SANTANDER CAFE  33,552.58          10,221.69
## 61  2009      PUTUMAYO CAFE      23           26.7
```



## 62	2009	QUINDIO	CAFE	19,052.00	21,985.00
## 63	2009	RISARALDA	CAFE	45,428.00	53,648.00
## 64	2009	SANTANDER	CAFE	37,985.90	26,311.61
## 65	2009	TOLIMA	CAFE	88,667.00	88,633.10
## 66	2009	VALLE DEL CAUCA	CAFE	67,001.30	62,711.08
##	Rendimiento..ha.ton. Produccion.Nacional..ton. Area.Nacional..ha.				
## 45				14.63	14.90
## 46				0.04	0.10
## 47				1.21	1.41
## 48				11.52	9.68
## 49				0.33	0.31
## 50				0.29	0.25
## 51				6.66	7.67
## 52				1.80	3.10
## 53				0.01	0.01
## 54				5.24	5.76
## 55				14.76	11.49
## 56				0.33	0.59
## 57				1.89	2.26
## 58				0.24	0.29
## 59				3.88	3.51
## 60				1.44	4.45
## 61				0.00	0.00
## 62				3.10	2.52
## 63				7.57	6.02
## 64				3.71	5.03
## 65				12.50	11.75
## 66				8.85	8.88

```
subset(cafe_df, subset = Anio == '2012')
```

##	Anio	Departamento	Producto	Area..ha.	Produccion..ton.
## 113	2012	ANTIOQUIA	CAFE	112,221.14	91,621.30
## 114	2012	BOLIVAR	CAFE	870	652.5
## 115	2012	BOYACA	CAFE	6,698.20	4,981.59
## 116	2012	CALDAS	CAFE	54,871.88	54,115.96
## 117	2012	CAQUETA	CAFE	2,882.50	2,446.38
## 118	2012	CASANARE	CAFE	2,322.00	1,718.25
## 119	2012	CAUCA	CAFE	56,825.00	50,588.14
## 120	2012	CESAR	CAFE	22,911.00	19,994.35
## 121	2012	CHOCO	CAFE	70	140
## 122	2012	CUNDINAMARCA	CAFE	37,175.06	30,786.41
## 123	2012	GUAVIARE	CAFE	0	0
## 124	2012	HUILA	CAFE	79,809.34	85,212.64
## 125	2012	LA GUAJIRA	CAFE	5,143.00	3,434.30
## 126	2012	MAGDALENA	CAFE	17,686.00	14,096.05
## 127	2012	META	CAFE	2,783.00	2,133.10
## 128	2012	NARIÑO	CAFE	27,806.40	28,077.94
## 129	2012	NORTE DE SANTANDER	CAFE	19,339.31	12,214.54
## 130	2012	PUTUMAYO	CAFE	42	48.4
## 131	2012	QUINDIO	CAFE	21,109.83	18,030.13
## 132	2012	RISARALDA	CAFE	45,588.03	36,989.43
## 133	2012	SANTANDER	CAFE	33,947.15	23,271.89
## 134	2012	TOLIMA	CAFE	90,904.48	85,027.49
## 135	2012	VALLE DEL CAUCA	CAFE	69,456.71	61,190.55

##	Rendimiento..ha.ton.	Produccion.Nacional..ton.	Area.Nacional..ha.
## 113	0.82	14.62	15.80
## 114	0.75	0.10	0.12
## 115	0.74	0.79	0.94
## 116	0.99	8.63	7.72
## 117	0.85	0.39	0.41
## 118	0.74	0.27	0.33
## 119	0.89	8.07	8.00
## 120	0.87	3.19	3.22
## 121	2.00	0.02	0.01
## 122	0.83	4.91	5.23
## 123	0.00	0.00	0.00
## 124	1.07	13.60	11.23
## 125	0.67	0.55	0.72
## 126	0.80	2.25	2.49
## 127	0.77	0.34	0.39
## 128	1.01	4.48	3.91
## 129	0.63	1.95	2.72
## 130	1.15	0.01	0.01
## 131	0.85	2.88	2.97
## 132	0.81	5.90	6.42
## 133	0.69	3.71	4.78
## 134	0.94	13.57	12.80
## 135	0.88	9.76	9.78

```
subset(cafe_df, subset=Anio == '2009',
       select=c('Rendimiento..ha.ton.', 'Produccion.Nacional..ton.'))
```

##	Rendimiento..ha.ton.	Produccion.Nacional..ton.
## 45	0.92	14.63
## 46	0.38	0.04
## 47	0.80	1.21
## 48	1.12	11.52
## 49	1.00	0.33
## 50	1.09	0.29
## 51	0.82	6.66
## 52	0.55	1.80
## 53	1.13	0.01
## 54	0.85	5.24
## 55	1.21	14.76
## 56	0.52	0.33
## 57	0.79	1.89
## 58	0.75	0.24
## 59	1.04	3.88
## 60	0.30	1.44
## 61	1.16	0.00
## 62	1.15	3.10
## 63	1.18	7.57
## 64	0.69	3.71
## 65	1.00	12.50
## 66	0.94	8.85

```
subset(cafe_df, subset=Departamento == 'CUNDINAMARCA',
       select=c('Rendimiento..ha.ton.', 'Produccion.Nacional..ton.'))
```

```
## Rendimiento..ha.ton. Produccion.Nacional..ton.
## 10 0.78 4.07
## 32 1.79 9.44
## 54 0.85 5.24
## 77 0.84 4.78
## 100 0.87 5.12
## 122 0.83 4.91
## 145 0.69 3.83
## 167 0.75 3.45
## 189 0.91 3.66
## 211 0.95 3.68
## 232 1.10 3.99
## 254 1.12 3.81
```

```
subset(caf _df, subset=Departamento == 'CUNDINAMARCA' & Anio <= '2012',
       select=c( 'Produccion.Nacional..ton.', 'Area.Nacional..ha.', 'Rendimiento..ha.ton.'))
```

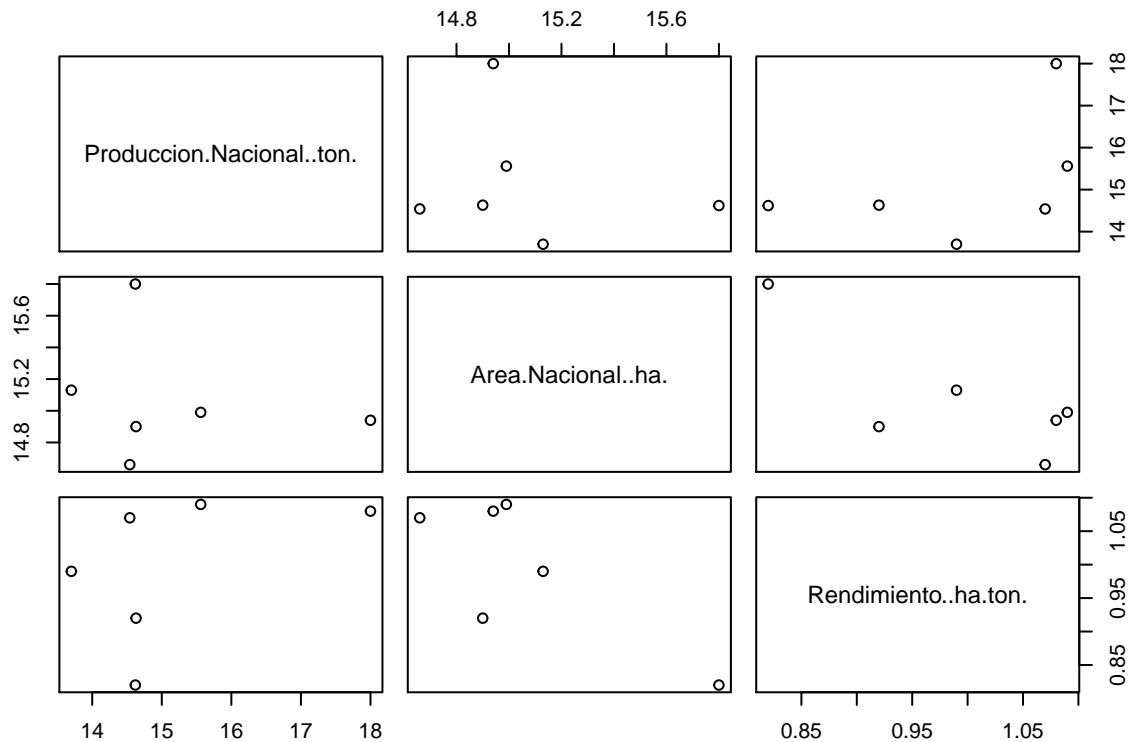
```
## Produccion.Nacional..ton. Area.Nacional..ha. Rendimiento..ha.ton.
## 10 4.07 5.61 0.78
## 32 9.44 5.75 1.79
## 54 5.24 5.76 0.85
## 77 4.78 5.95 0.84
## 100 5.12 5.26 0.87
## 122 4.91 5.23 0.83
```

#PONEMOS LA INSTRUCCION EN UNA VARIABLE Y LO GRAFICAMOS

```
Grupo_Antioquia<-subset(caf _df, subset=Departamento == 'ANTIOQUIA' & Anio <= '2012',
                        select=c( 'Produccion.Nacional..ton.', 'Area.Nacional..ha.', 'Rendimiento..ha.ton.'))
Grupo_Antioquia
```

```
## Produccion.Nacional..ton. Area.Nacional..ha. Rendimiento..ha.ton.
## 1 14.54 14.66 1.07
## 23 13.70 15.13 0.99
## 45 14.63 14.90 0.92
## 67 15.56 14.99 1.09
## 90 18.00 14.94 1.08
## 113 14.62 15.80 0.82
```

```
plot(Grupo_Antioquia)
```



#PODEMOS VER SI EXISTEN DATOS NULOS EN EL DATAFRAME

```
is.na(cafe_df)
```

```
##      Anio Departamento Producto Area..ha. Produccion..ton.
## [1,] FALSE           FALSE    FALSE    FALSE           FALSE
## [2,] FALSE           FALSE    FALSE    FALSE           FALSE
## [3,] FALSE           FALSE    FALSE    FALSE           FALSE
## [4,] FALSE           FALSE    FALSE    FALSE           FALSE
## [5,] FALSE           FALSE    FALSE    FALSE           FALSE
## [6,] FALSE           FALSE    FALSE    FALSE           FALSE
## [7,] FALSE           FALSE    FALSE    FALSE           FALSE
## [8,] FALSE           FALSE    FALSE    FALSE           FALSE
## [9,] FALSE           FALSE    FALSE    FALSE           FALSE
## [10,] FALSE          FALSE    FALSE    FALSE           FALSE
## [11,] FALSE          FALSE    FALSE    FALSE           FALSE
## [12,] FALSE          FALSE    FALSE    FALSE           FALSE
## [13,] FALSE          FALSE    FALSE    FALSE           FALSE
## [14,] FALSE          FALSE    FALSE    FALSE           FALSE
## [15,] FALSE          FALSE    FALSE    FALSE           FALSE
## [16,] FALSE          FALSE    FALSE    FALSE           FALSE
## [17,] FALSE          FALSE    FALSE    FALSE           FALSE
## [18,] FALSE          FALSE    FALSE    FALSE           FALSE
## [19,] FALSE          FALSE    FALSE    FALSE           FALSE
## [20,] FALSE          FALSE    FALSE    FALSE           FALSE
## [21,] FALSE          FALSE    FALSE    FALSE           FALSE
## [22,] FALSE          FALSE    FALSE    FALSE           FALSE
## [23,] FALSE          FALSE    FALSE    FALSE           FALSE
## [24,] FALSE          FALSE    FALSE    FALSE           FALSE
## [25,] FALSE          FALSE    FALSE    FALSE           FALSE
## [26,] FALSE          FALSE    FALSE    FALSE           FALSE
```

[illegible]

[illegible]

[illegible]

[illegible]



##	[243,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[244,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[245,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[246,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[247,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[248,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[249,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[250,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[251,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[252,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[253,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[254,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[255,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[256,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[257,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[258,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[259,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[260,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[261,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[262,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[263,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[264,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[265,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	[266,]	FALSE	FALSE	FALSE	FALSE	FALSE
##	Rendimiento..ha.ton. Produccion.Nacional..ton. Area.Nacional..ha.					
##	[1,]	FALSE		FALSE		FALSE
##	[2,]	FALSE		FALSE		FALSE
##	[3,]	FALSE		FALSE		FALSE
##	[4,]	FALSE		FALSE		FALSE
##	[5,]	FALSE		FALSE		FALSE
##	[6,]	FALSE		FALSE		FALSE
##	[7,]	FALSE		FALSE		FALSE
##	[8,]	FALSE		FALSE		FALSE
##	[9,]	FALSE		FALSE		FALSE
##	[10,]	FALSE		FALSE		FALSE
##	[11,]	FALSE		FALSE		FALSE
##	[12,]	FALSE		FALSE		FALSE
##	[13,]	FALSE		FALSE		FALSE
##	[14,]	FALSE		FALSE		FALSE
##	[15,]	FALSE		FALSE		FALSE
##	[16,]	FALSE		FALSE		FALSE
##	[17,]	FALSE		FALSE		FALSE
##	[18,]	FALSE		FALSE		FALSE
##	[19,]	FALSE		FALSE		FALSE
##	[20,]	FALSE		FALSE		FALSE
##	[21,]	FALSE		FALSE		FALSE
##	[22,]	FALSE		FALSE		FALSE
##	[23,]	FALSE		FALSE		FALSE
##	[24,]	FALSE		FALSE		FALSE
##	[25,]	FALSE		FALSE		FALSE
##	[26,]	FALSE		FALSE		FALSE
##	[27,]	FALSE		FALSE		FALSE
##	[28,]	FALSE		FALSE		FALSE
##	[29,]	FALSE		FALSE		FALSE

##	[30,]	FALSE	FALSE	FALSE
##	[31,]	FALSE	FALSE	FALSE
##	[32,]	FALSE	FALSE	FALSE
##	[33,]	FALSE	FALSE	FALSE
##	[34,]	FALSE	FALSE	FALSE
##	[35,]	FALSE	FALSE	FALSE
##	[36,]	FALSE	FALSE	FALSE
##	[37,]	FALSE	FALSE	FALSE
##	[38,]	FALSE	FALSE	FALSE
##	[39,]	FALSE	FALSE	FALSE
##	[40,]	FALSE	FALSE	FALSE
##	[41,]	FALSE	FALSE	FALSE
##	[42,]	FALSE	FALSE	FALSE
##	[43,]	FALSE	FALSE	FALSE
##	[44,]	FALSE	FALSE	FALSE
##	[45,]	FALSE	FALSE	FALSE
##	[46,]	FALSE	FALSE	FALSE
##	[47,]	FALSE	FALSE	FALSE
##	[48,]	FALSE	FALSE	FALSE
##	[49,]	FALSE	FALSE	FALSE
##	[50,]	FALSE	FALSE	FALSE
##	[51,]	FALSE	FALSE	FALSE
##	[52,]	FALSE	FALSE	FALSE
##	[53,]	FALSE	FALSE	FALSE
##	[54,]	FALSE	FALSE	FALSE
##	[55,]	FALSE	FALSE	FALSE
##	[56,]	FALSE	FALSE	FALSE
##	[57,]	FALSE	FALSE	FALSE
##	[58,]	FALSE	FALSE	FALSE
##	[59,]	FALSE	FALSE	FALSE
##	[60,]	FALSE	FALSE	FALSE
##	[61,]	FALSE	FALSE	FALSE
##	[62,]	FALSE	FALSE	FALSE
##	[63,]	FALSE	FALSE	FALSE
##	[64,]	FALSE	FALSE	FALSE
##	[65,]	FALSE	FALSE	FALSE
##	[66,]	FALSE	FALSE	FALSE
##	[67,]	FALSE	FALSE	FALSE
##	[68,]	FALSE	FALSE	FALSE
##	[69,]	FALSE	FALSE	FALSE
##	[70,]	FALSE	FALSE	FALSE
##	[71,]	FALSE	FALSE	FALSE
##	[72,]	FALSE	FALSE	FALSE
##	[73,]	FALSE	FALSE	FALSE
##	[74,]	FALSE	FALSE	FALSE
##	[75,]	FALSE	FALSE	FALSE
##	[76,]	FALSE	FALSE	FALSE
##	[77,]	FALSE	FALSE	FALSE
##	[78,]	FALSE	FALSE	FALSE
##	[79,]	FALSE	FALSE	FALSE
##	[80,]	FALSE	FALSE	FALSE
##	[81,]	FALSE	FALSE	FALSE
##	[82,]	FALSE	FALSE	FALSE
##	[83,]	FALSE	FALSE	FALSE

## [84,]	FALSE	FALSE	FALSE
## [85,]	FALSE	FALSE	FALSE
## [86,]	FALSE	FALSE	FALSE
## [87,]	FALSE	FALSE	FALSE
## [88,]	FALSE	FALSE	FALSE
## [89,]	FALSE	FALSE	FALSE
## [90,]	FALSE	FALSE	FALSE
## [91,]	FALSE	FALSE	FALSE
## [92,]	FALSE	FALSE	FALSE
## [93,]	FALSE	FALSE	FALSE
## [94,]	FALSE	FALSE	FALSE
## [95,]	FALSE	FALSE	FALSE
## [96,]	FALSE	FALSE	FALSE
## [97,]	FALSE	FALSE	FALSE
## [98,]	FALSE	FALSE	FALSE
## [99,]	FALSE	FALSE	FALSE
## [100,]	FALSE	FALSE	FALSE
## [101,]	FALSE	FALSE	FALSE
## [102,]	FALSE	FALSE	FALSE
## [103,]	FALSE	FALSE	FALSE
## [104,]	FALSE	FALSE	FALSE
## [105,]	FALSE	FALSE	FALSE
## [106,]	FALSE	FALSE	FALSE
## [107,]	FALSE	FALSE	FALSE
## [108,]	FALSE	FALSE	FALSE
## [109,]	FALSE	FALSE	FALSE
## [110,]	FALSE	FALSE	FALSE
## [111,]	FALSE	FALSE	FALSE
## [112,]	FALSE	FALSE	FALSE
## [113,]	FALSE	FALSE	FALSE
## [114,]	FALSE	FALSE	FALSE
## [115,]	FALSE	FALSE	FALSE
## [116,]	FALSE	FALSE	FALSE
## [117,]	FALSE	FALSE	FALSE
## [118,]	FALSE	FALSE	FALSE
## [119,]	FALSE	FALSE	FALSE
## [120,]	FALSE	FALSE	FALSE
## [121,]	FALSE	FALSE	FALSE
## [122,]	FALSE	FALSE	FALSE
## [123,]	FALSE	FALSE	FALSE
## [124,]	FALSE	FALSE	FALSE
## [125,]	FALSE	FALSE	FALSE
## [126,]	FALSE	FALSE	FALSE
## [127,]	FALSE	FALSE	FALSE
## [128,]	FALSE	FALSE	FALSE
## [129,]	FALSE	FALSE	FALSE
## [130,]	FALSE	FALSE	FALSE
## [131,]	FALSE	FALSE	FALSE
## [132,]	FALSE	FALSE	FALSE
## [133,]	FALSE	FALSE	FALSE
## [134,]	FALSE	FALSE	FALSE
## [135,]	FALSE	FALSE	FALSE
## [136,]	FALSE	FALSE	FALSE
## [137,]	FALSE	FALSE	FALSE

## [138,]	FALSE	FALSE	FALSE
## [139,]	FALSE	FALSE	FALSE
## [140,]	FALSE	FALSE	FALSE
## [141,]	FALSE	FALSE	FALSE
## [142,]	FALSE	FALSE	FALSE
## [143,]	FALSE	FALSE	FALSE
## [144,]	FALSE	FALSE	FALSE
## [145,]	FALSE	FALSE	FALSE
## [146,]	FALSE	FALSE	FALSE
## [147,]	FALSE	FALSE	FALSE
## [148,]	FALSE	FALSE	FALSE
## [149,]	FALSE	FALSE	FALSE
## [150,]	FALSE	FALSE	FALSE
## [151,]	FALSE	FALSE	FALSE
## [152,]	FALSE	FALSE	FALSE
## [153,]	FALSE	FALSE	FALSE
## [154,]	FALSE	FALSE	FALSE
## [155,]	FALSE	FALSE	FALSE
## [156,]	FALSE	FALSE	FALSE
## [157,]	FALSE	FALSE	FALSE
## [158,]	FALSE	FALSE	FALSE
## [159,]	FALSE	FALSE	FALSE
## [160,]	FALSE	FALSE	FALSE
## [161,]	FALSE	FALSE	FALSE
## [162,]	FALSE	FALSE	FALSE
## [163,]	FALSE	FALSE	FALSE
## [164,]	FALSE	FALSE	FALSE
## [165,]	FALSE	FALSE	FALSE
## [166,]	FALSE	FALSE	FALSE
## [167,]	FALSE	FALSE	FALSE
## [168,]	FALSE	FALSE	FALSE
## [169,]	FALSE	FALSE	FALSE
## [170,]	FALSE	FALSE	FALSE
## [171,]	FALSE	FALSE	FALSE
## [172,]	FALSE	FALSE	FALSE
## [173,]	FALSE	FALSE	FALSE
## [174,]	FALSE	FALSE	FALSE
## [175,]	FALSE	FALSE	FALSE
## [176,]	FALSE	FALSE	FALSE
## [177,]	FALSE	FALSE	FALSE
## [178,]	FALSE	FALSE	FALSE
## [179,]	FALSE	FALSE	FALSE
## [180,]	FALSE	FALSE	FALSE
## [181,]	FALSE	FALSE	FALSE
## [182,]	FALSE	FALSE	FALSE
## [183,]	FALSE	FALSE	FALSE
## [184,]	FALSE	FALSE	FALSE
## [185,]	FALSE	FALSE	FALSE
## [186,]	FALSE	FALSE	FALSE
## [187,]	FALSE	FALSE	FALSE
## [188,]	FALSE	FALSE	FALSE
## [189,]	FALSE	FALSE	FALSE
## [190,]	FALSE	FALSE	FALSE
## [191,]	FALSE	FALSE	FALSE

## [192,]	FALSE	FALSE	FALSE
## [193,]	FALSE	FALSE	FALSE
## [194,]	FALSE	FALSE	FALSE
## [195,]	FALSE	FALSE	FALSE
## [196,]	FALSE	FALSE	FALSE
## [197,]	FALSE	FALSE	FALSE
## [198,]	FALSE	FALSE	FALSE
## [199,]	FALSE	FALSE	FALSE
## [200,]	FALSE	FALSE	FALSE
## [201,]	FALSE	FALSE	FALSE
## [202,]	FALSE	FALSE	FALSE
## [203,]	FALSE	FALSE	FALSE
## [204,]	FALSE	FALSE	FALSE
## [205,]	FALSE	FALSE	FALSE
## [206,]	FALSE	FALSE	FALSE
## [207,]	FALSE	FALSE	FALSE
## [208,]	FALSE	FALSE	FALSE
## [209,]	FALSE	FALSE	FALSE
## [210,]	FALSE	FALSE	FALSE
## [211,]	FALSE	FALSE	FALSE
## [212,]	FALSE	FALSE	FALSE
## [213,]	FALSE	FALSE	FALSE
## [214,]	FALSE	FALSE	FALSE
## [215,]	FALSE	FALSE	FALSE
## [216,]	FALSE	FALSE	FALSE
## [217,]	FALSE	FALSE	FALSE
## [218,]	FALSE	FALSE	FALSE
## [219,]	FALSE	FALSE	FALSE
## [220,]	FALSE	FALSE	FALSE
## [221,]	FALSE	FALSE	FALSE
## [222,]	FALSE	FALSE	FALSE
## [223,]	FALSE	FALSE	FALSE
## [224,]	FALSE	FALSE	FALSE
## [225,]	FALSE	FALSE	FALSE
## [226,]	FALSE	FALSE	FALSE
## [227,]	FALSE	FALSE	FALSE
## [228,]	FALSE	FALSE	FALSE
## [229,]	FALSE	FALSE	FALSE
## [230,]	FALSE	FALSE	FALSE
## [231,]	FALSE	FALSE	FALSE
## [232,]	FALSE	FALSE	FALSE
## [233,]	FALSE	FALSE	FALSE
## [234,]	FALSE	FALSE	FALSE
## [235,]	FALSE	FALSE	FALSE
## [236,]	FALSE	FALSE	FALSE
## [237,]	FALSE	FALSE	FALSE
## [238,]	FALSE	FALSE	FALSE
## [239,]	FALSE	FALSE	FALSE
## [240,]	FALSE	FALSE	FALSE
## [241,]	FALSE	FALSE	FALSE
## [242,]	FALSE	FALSE	FALSE
## [243,]	FALSE	FALSE	FALSE
## [244,]	FALSE	FALSE	FALSE
## [245,]	FALSE	FALSE	FALSE

##	[246,]	FALSE	FALSE	FALSE
##	[247,]	FALSE	FALSE	FALSE
##	[248,]	FALSE	FALSE	FALSE
##	[249,]	FALSE	FALSE	FALSE
##	[250,]	FALSE	FALSE	FALSE
##	[251,]	FALSE	FALSE	FALSE
##	[252,]	FALSE	FALSE	FALSE
##	[253,]	FALSE	FALSE	FALSE
##	[254,]	FALSE	FALSE	FALSE
##	[255,]	FALSE	FALSE	FALSE
##	[256,]	FALSE	FALSE	FALSE
##	[257,]	FALSE	FALSE	FALSE
##	[258,]	FALSE	FALSE	FALSE
##	[259,]	FALSE	FALSE	FALSE
##	[260,]	FALSE	FALSE	FALSE
##	[261,]	FALSE	FALSE	FALSE
##	[262,]	FALSE	FALSE	FALSE
##	[263,]	FALSE	FALSE	FALSE
##	[264,]	FALSE	FALSE	FALSE
##	[265,]	FALSE	FALSE	FALSE
##	[266,]	FALSE	FALSE	FALSE

```
is.na(cafe_df$Area..ha.)
```

[illegible]

```
is.na(caffe_df$Rendimiento..ha.ton.)
```

```
## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
```

```
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [253] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [265] FALSE FALSE
```

#NOS MUESTRA LA SUMA DE LOS VALORES NULOS EN EL DATAFRAME

```
rowSums(is.na(cafe_df))
```

```
## [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [38] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [75] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [112] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [149] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [186] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [223] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [260] 0 0 0 0 0 0 0
```

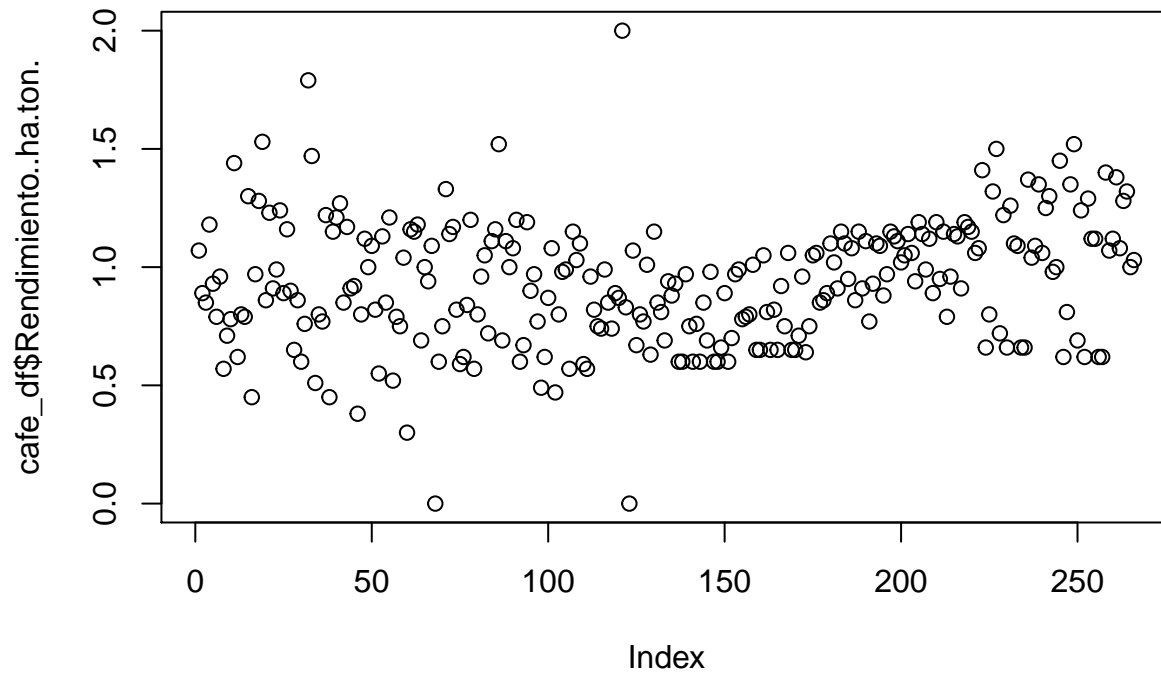
#NOS MUESTRA LA SUMA TOTAL DE TODOS LOS VALORES NULOS

```
sum(rowSums(is.na(cafe_df)))
```

```
## [1] 0
```

#VISUALIZAMOS HISTOGRAMAS POR VARIABLE

```
plot(cafe_df$Rendimiento..ha.ton.)
```

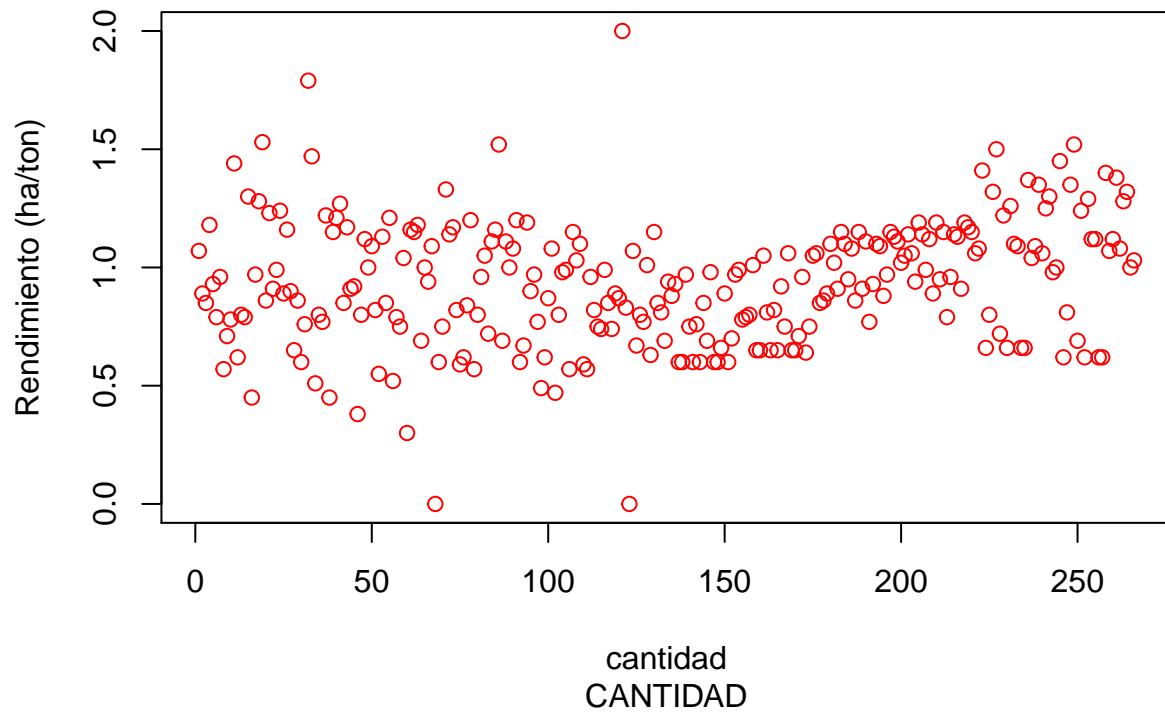


#GRAFICAMOS LA VARIABLE Y SELECCIONAMOS ALGUNOS DETALLES

```
plot(cafe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "CANTIDAD",  
     type = "p", # p indica puntos de dispersion  
     col = "red",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

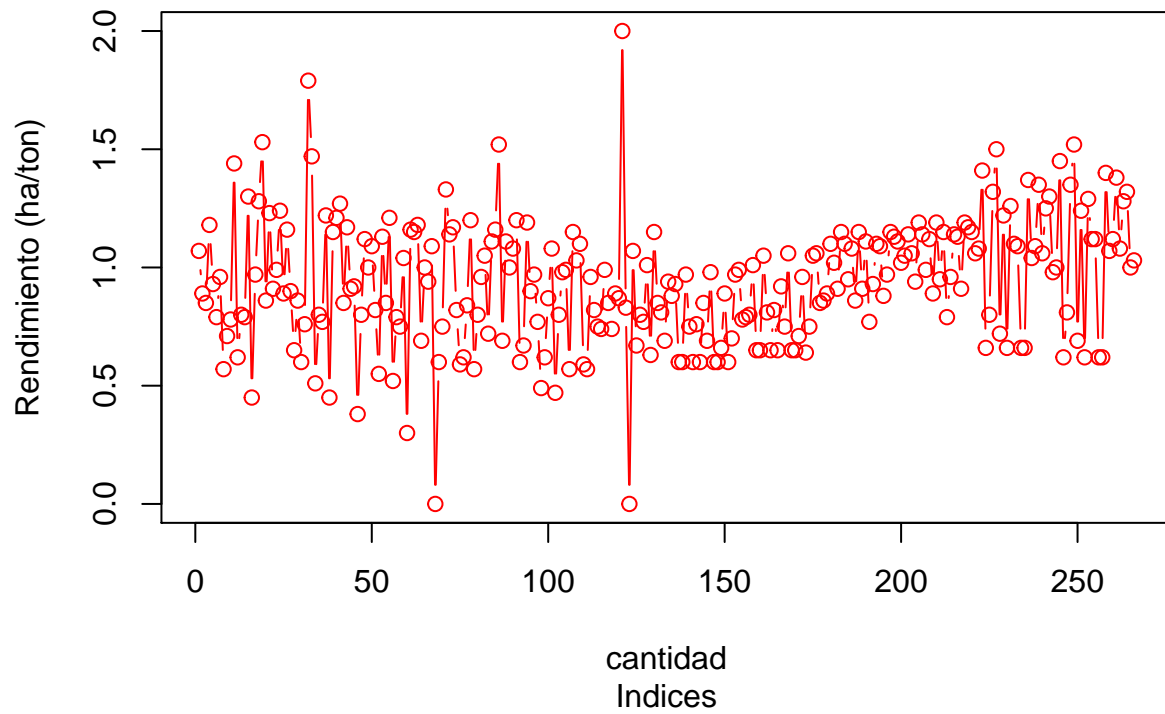


## Gráfico Dispersion



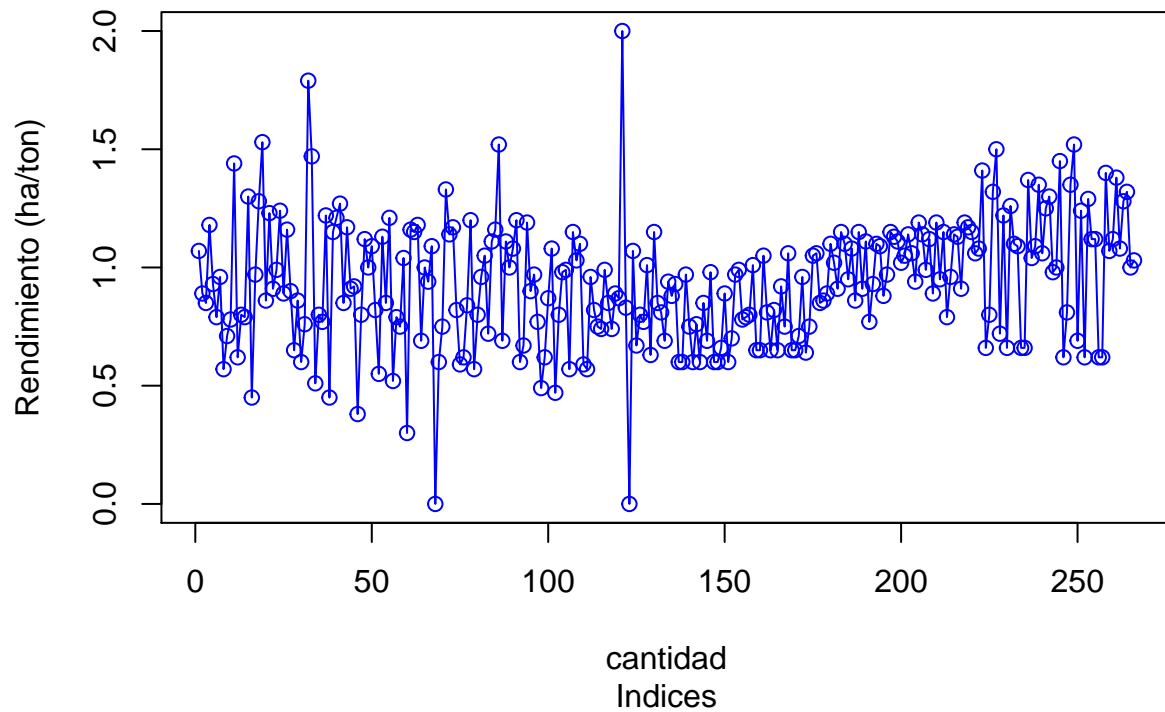
```
plot(caffe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "Indices",  
     type = "b", # b indica líneas uniendo los puntos de dispersion  
     col = "red",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

## Gráfico Dispersion



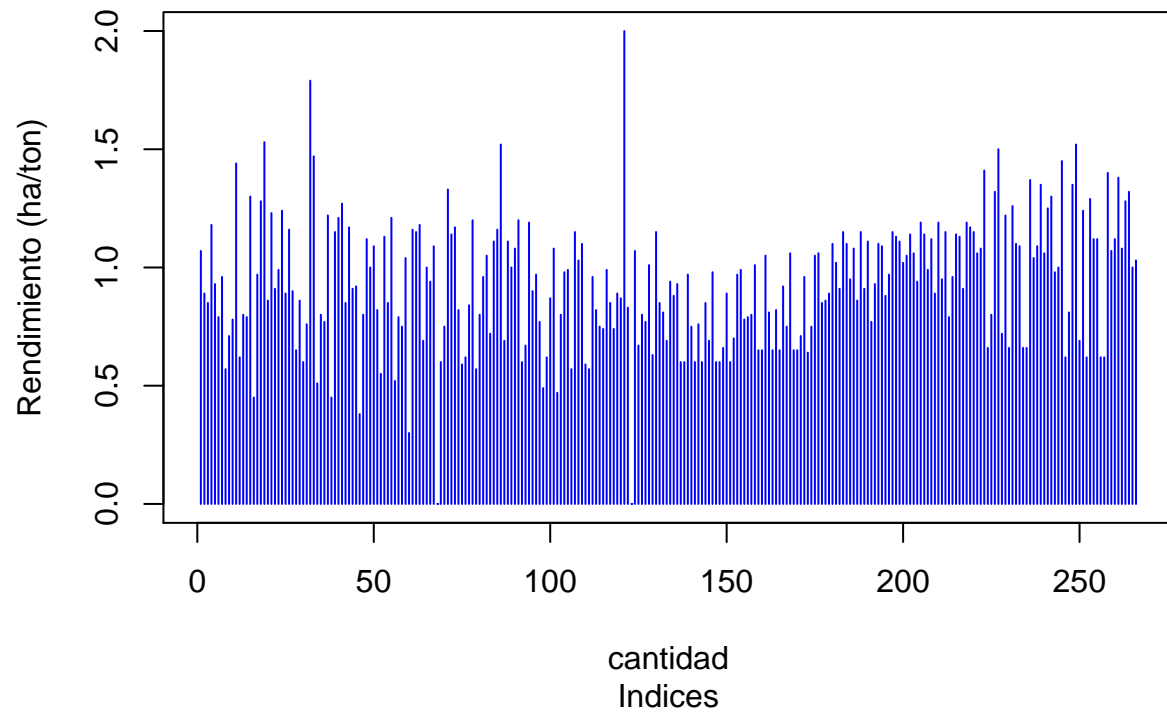
```
plot(caffe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "Indices",  
     type = "o", # b indica líneas sobreindicadas sobre los puntos de dispersion  
     col = "blue",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

## Gráfico Dispersion



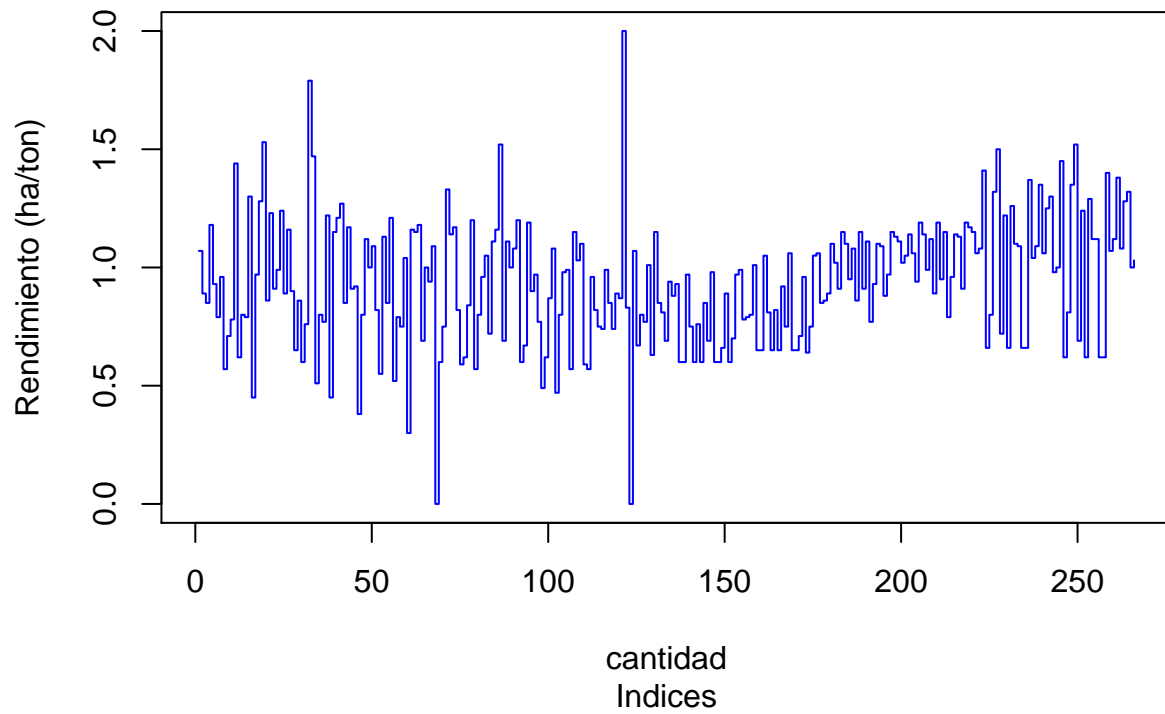
```
plot(caffe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "Indices",  
     type = "h", # b indica líneas tipo histograma sobre los puntos de dispersion  
     col = "blue",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

## Gráfico Dispersion



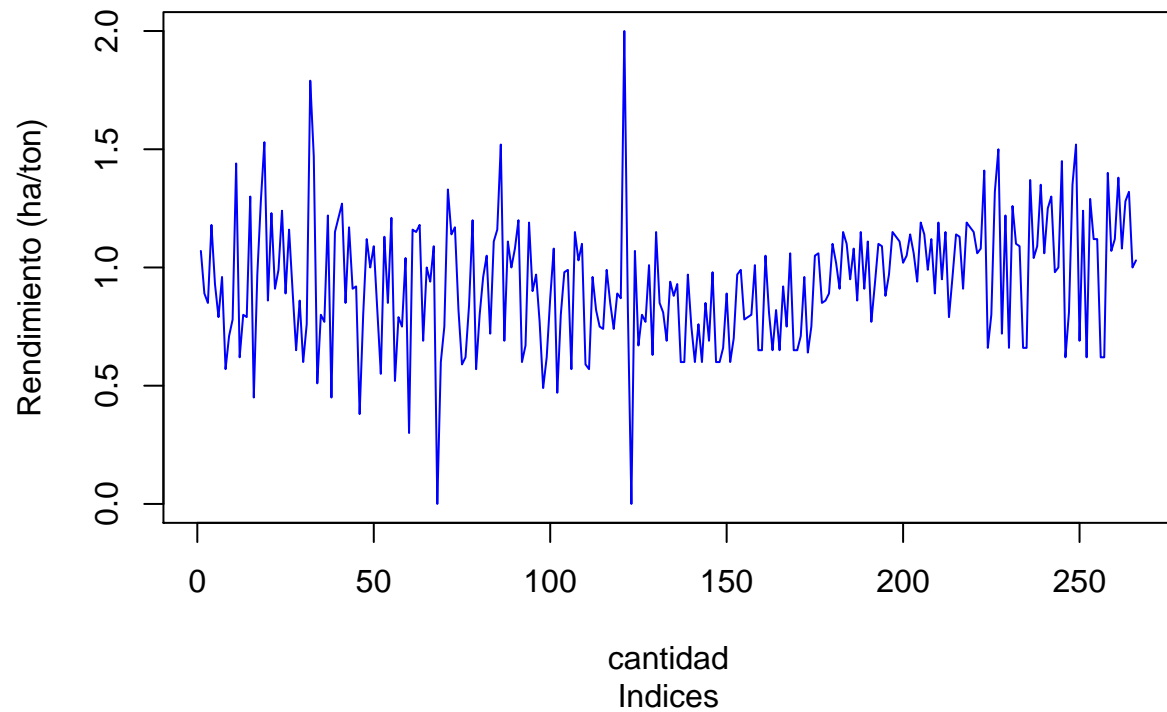
```
plot(caffe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "Indices",  
     type = "s", # s funcion escalera (horizontal a vertical)  
     col = "blue",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

## Gráfico Dispersion



```
plot(caffe_df$Rendimiento..ha.ton., main = "Gráfico Dispersion",  
     sub = "Indices",  
     type = "l", # l indica líneas  
     col = "blue",  
     xlab = "cantidad",  
     ylab = "Rendimiento (ha/ton)")
```

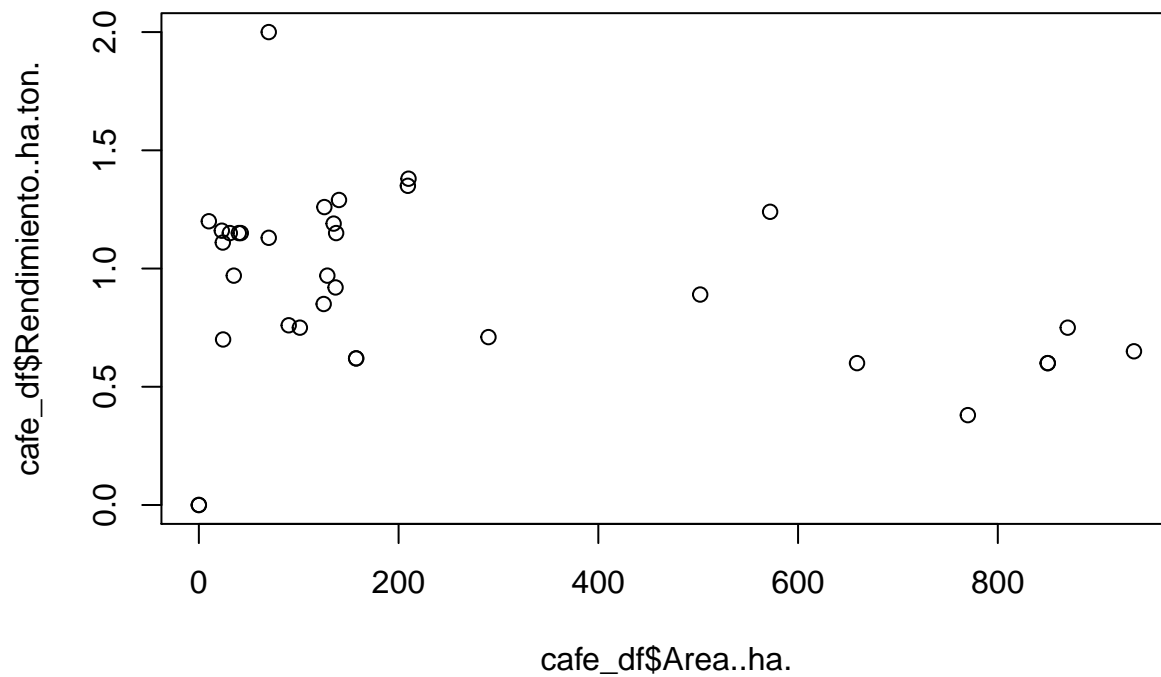
## Gráfico Dispersion



#GRAFICAS POR VARIABLES

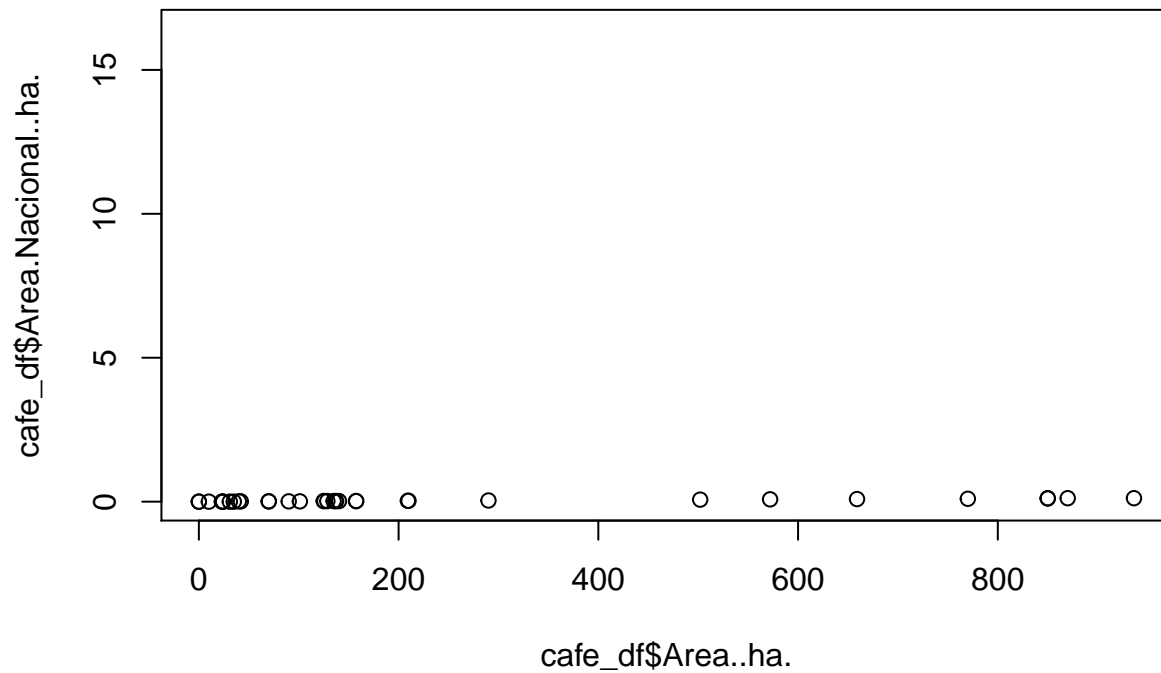
```
plot(x = cafe_df$Area..ha., y = cafe_df$Rendimiento..ha.ton.)
```

## Warning in xy.coords(x, y, xlabel, ylabel, log): NAs introduced by coercion

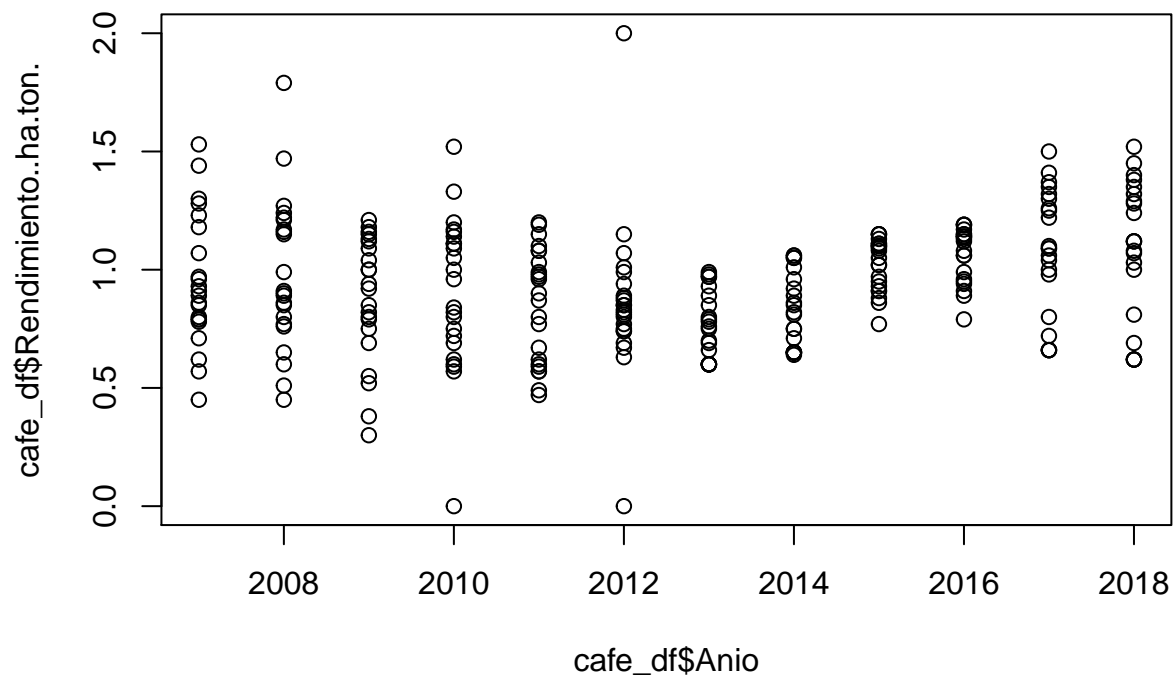


```
plot(x = cafe_df$Area..ha., y = cafe_df$Area.Nacional..ha.)
```

```
## Warning in xy.coords(x, y, xlabel, ylabel, log): NAs introduced by coercion
```

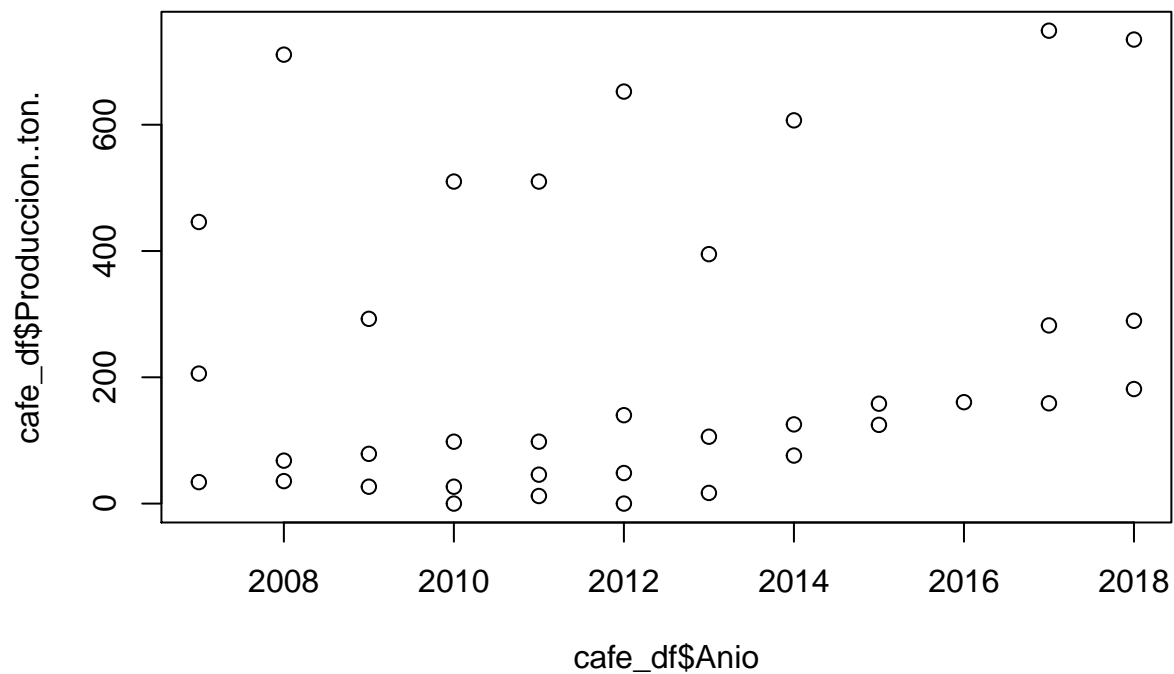


```
plot(x = cafe_df$Anio, y = cafe_df$Rendimiento..ha.ton.)
```



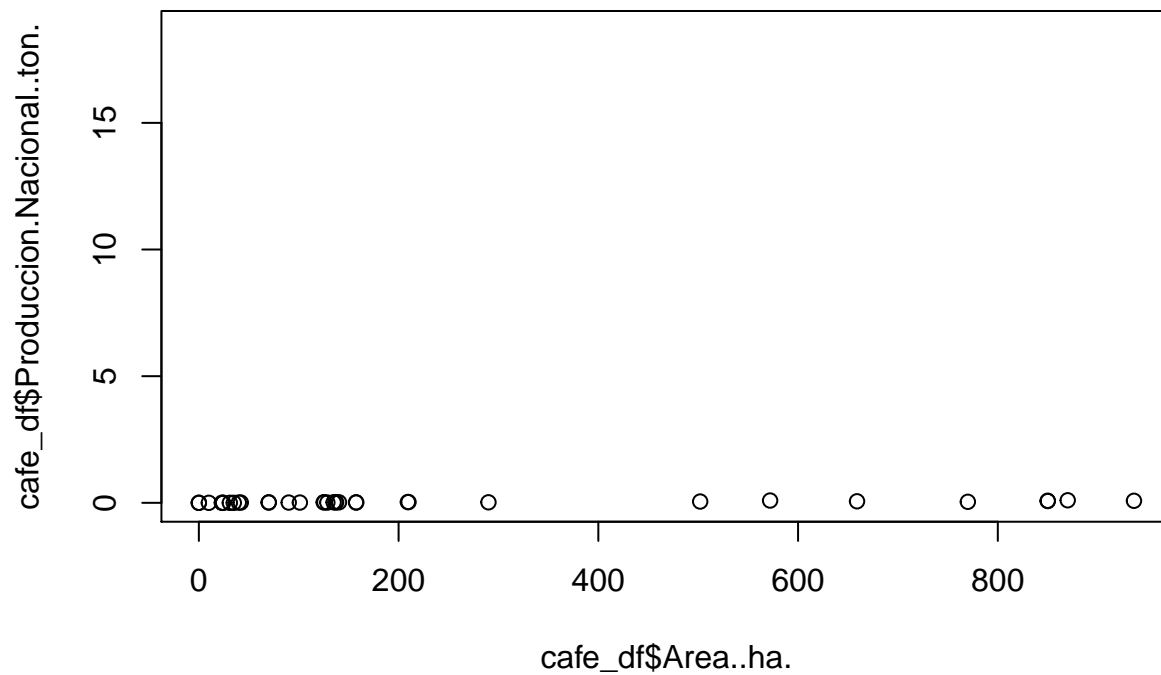
```
plot(x = cafe_df$Anio, y = cafe_df$Produccion..ton.)
```

```
## Warning in xy.coords(x, y, xlabel, ylabel, log): NAs introduced by coercion
```



```
plot(x = cafe_df$Area..ha., y = cafe_df$Produccion.Nacional..ton.)
```

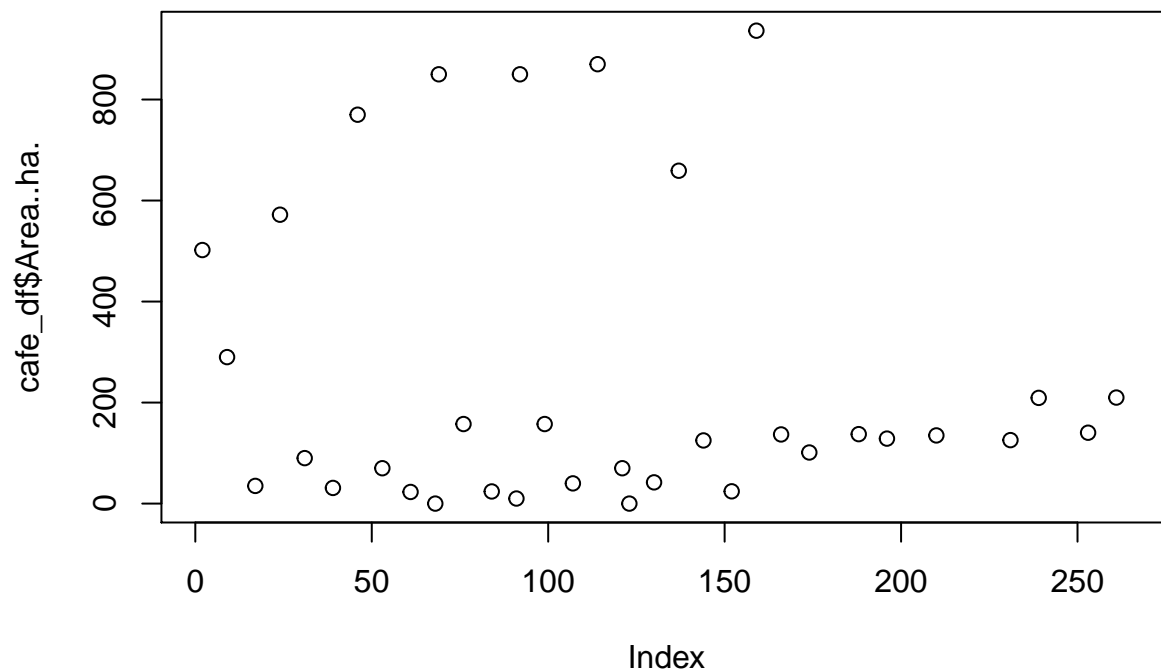
```
## Warning in xy.coords(x, y, xlabel, ylabel, log): NAs introduced by coercion
```



```
plot(cafe_df$Area..ha.)
```

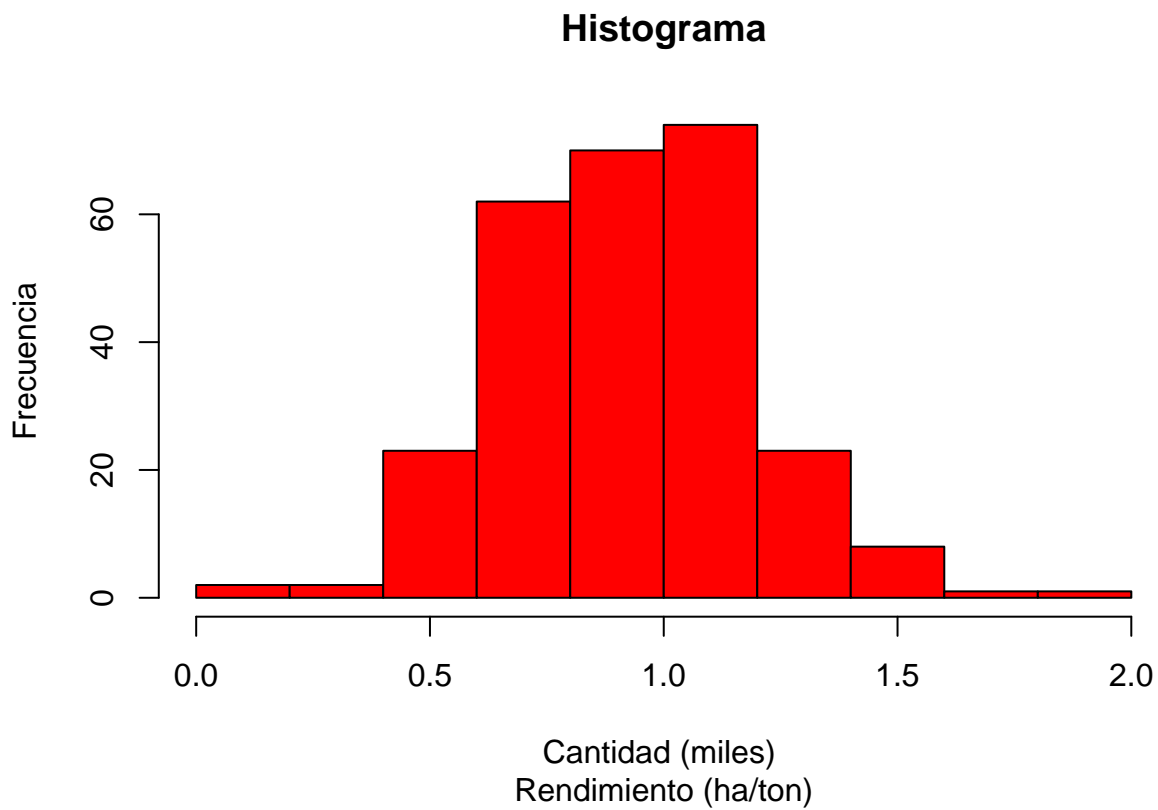
```
## Warning in xy.coords(x, y, xlabel, ylabel, log): NAs introduced by coercion
```





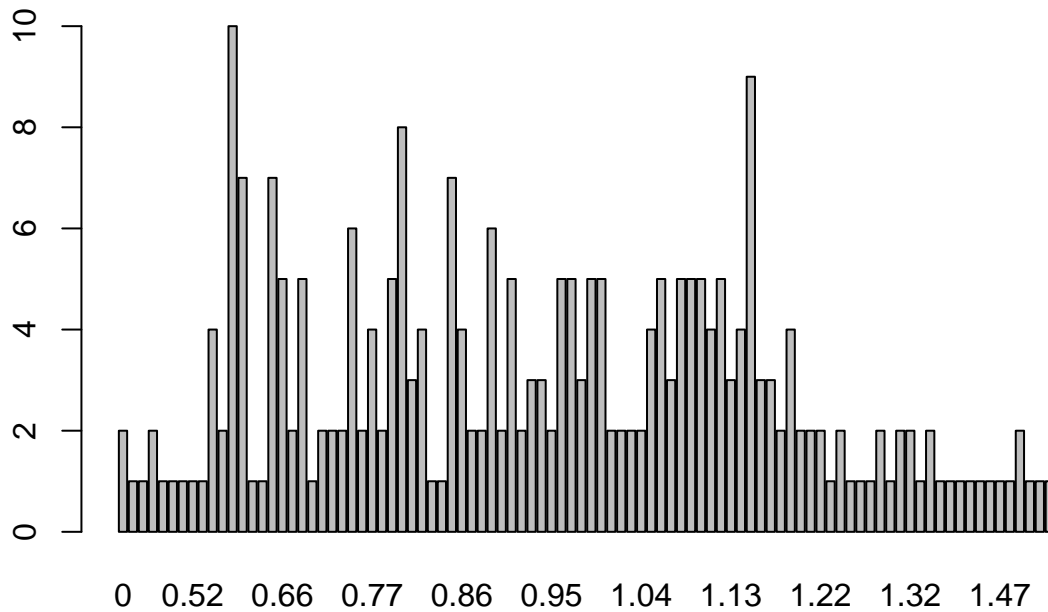
#REPRESENTA LA FRECUENCIA DE LOS DATOS EN UN HISTOGRAMA

```
hist(cafe_df$Rendimiento..ha.ton., main = "Histograma",
     sub = "Rendimiento (ha/ton)",
     col = "red",
     xlab = "Cantidad (miles) ",
     ylab = "Frecuencia")
```



#DIAGRAMA DE BARRAS DE LA COLUMNA RENDIMIENTO

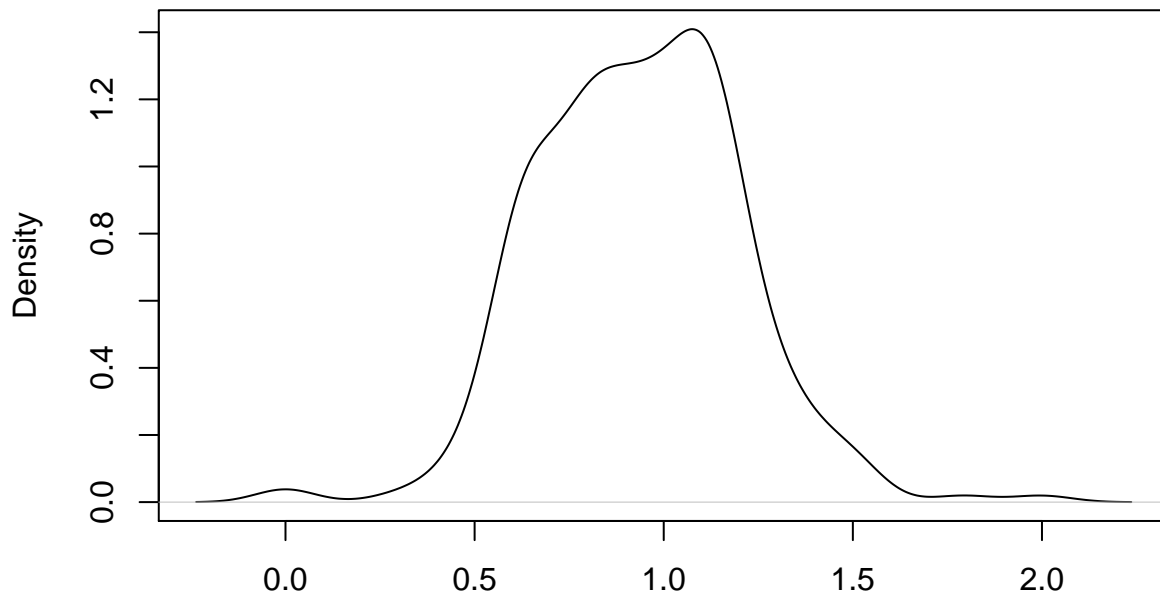
```
barplot(table(cafe_df$Rendimiento..ha.ton.))
```



#PARA VER COMO SE DISTRIBUYEN LOS DATOS PODEMOS USAR LA FUNCION DENSITY, LA DENSIDAD ES UNA VERSION SUAVIZADA DEL HISTOGRAMA, NOS PERMITE OBSERVAR SI LOS DATOS SE COMPORTAN COMO UNA DISTRIBUCION NORMAL

```
plot(density(cafe_df$Rendimiento..ha.ton.), main="Densidad para el Rendimiento (ha/ton)")
```

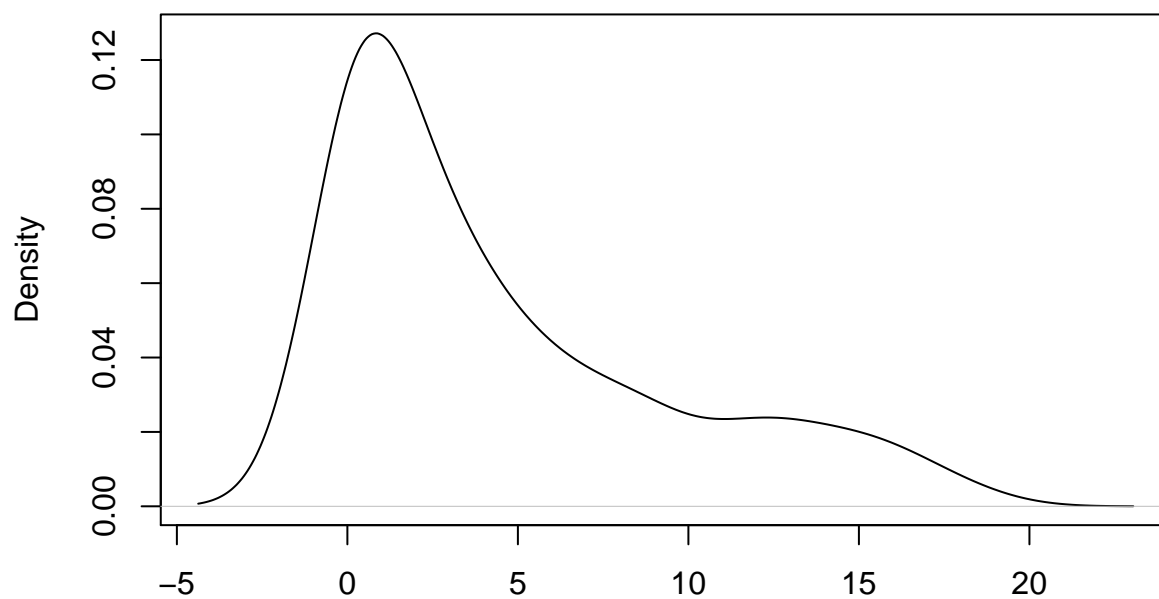
### Densidad para el Rendimiento (ha/ton)



N = 266 Bandwidth = 0.0787

```
plot(density(cafe_df$Produccion.Nacional..ton.), main="Densidad para la produccion Nacional (ton)")
```

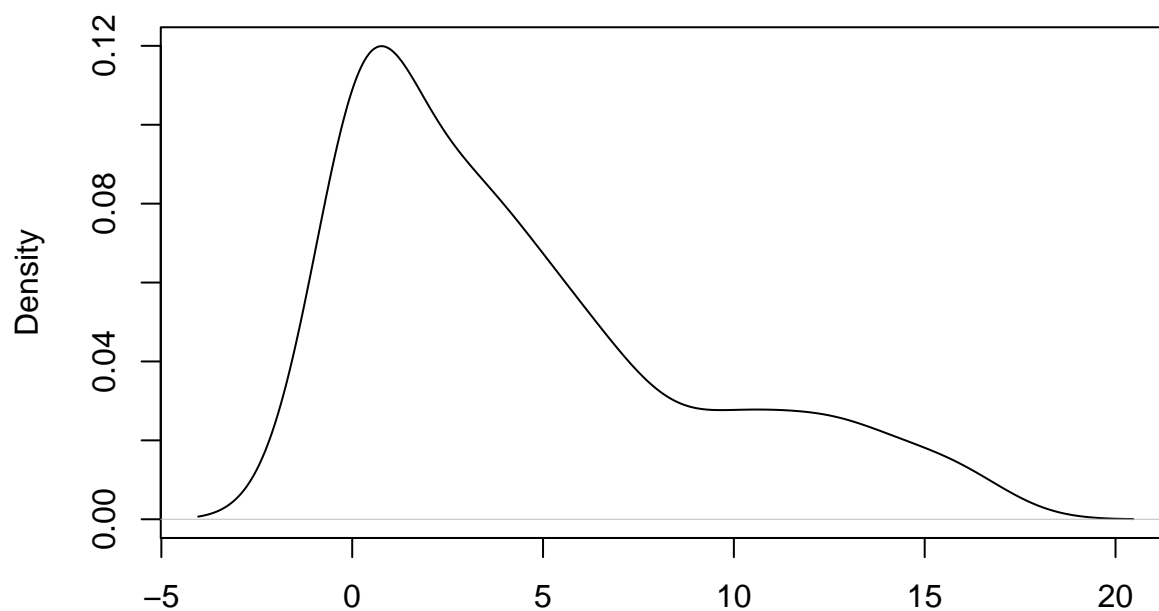
### Densidad para la produccion Nacional (ton)



N = 266 Bandwidth = 1.459

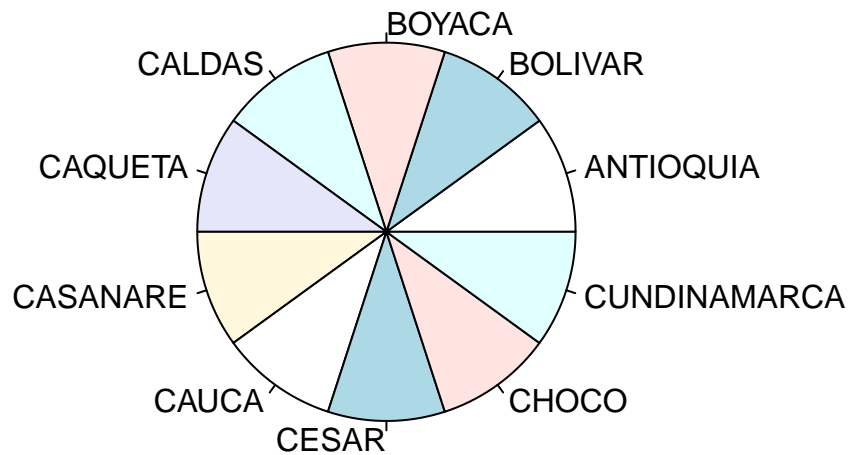
```
plot(density(cafe_df$Area.Nacional..ha.), main="Densidad para el Area Nacional (ha)")
```

### Densidad para el Area Nacional (ha)

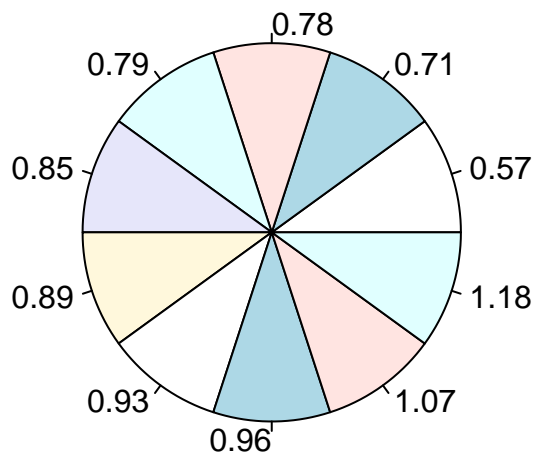


N = 266 Bandwidth = 1.345

```
#HACEMOS UN GRAFICO DE PASTEL  
pie(table(cafe_df$Departamento[1:10]))
```

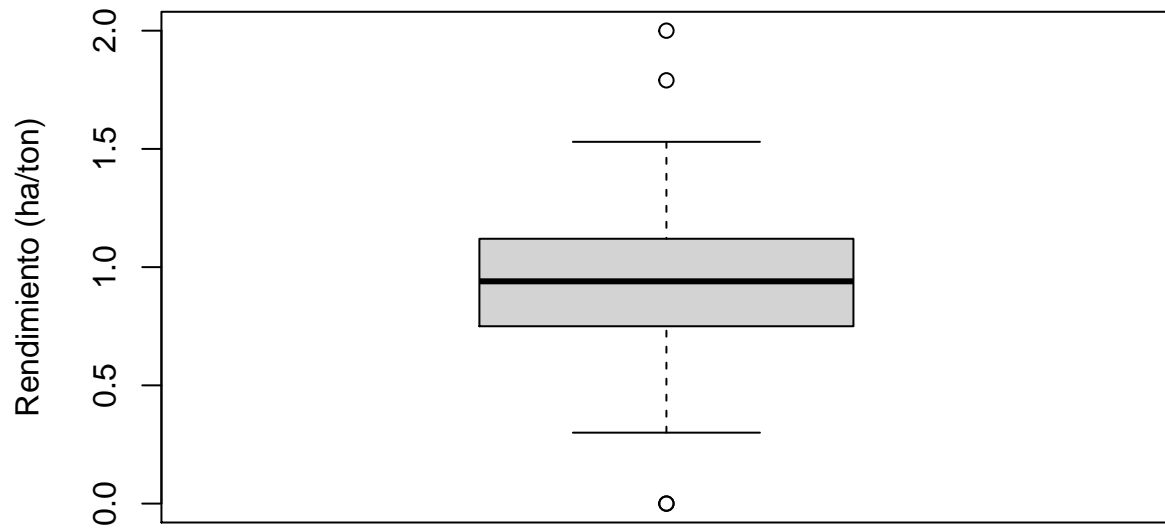


```
pie(table(cafe_df$Rendimiento..ha.ton.[1:10]))
```

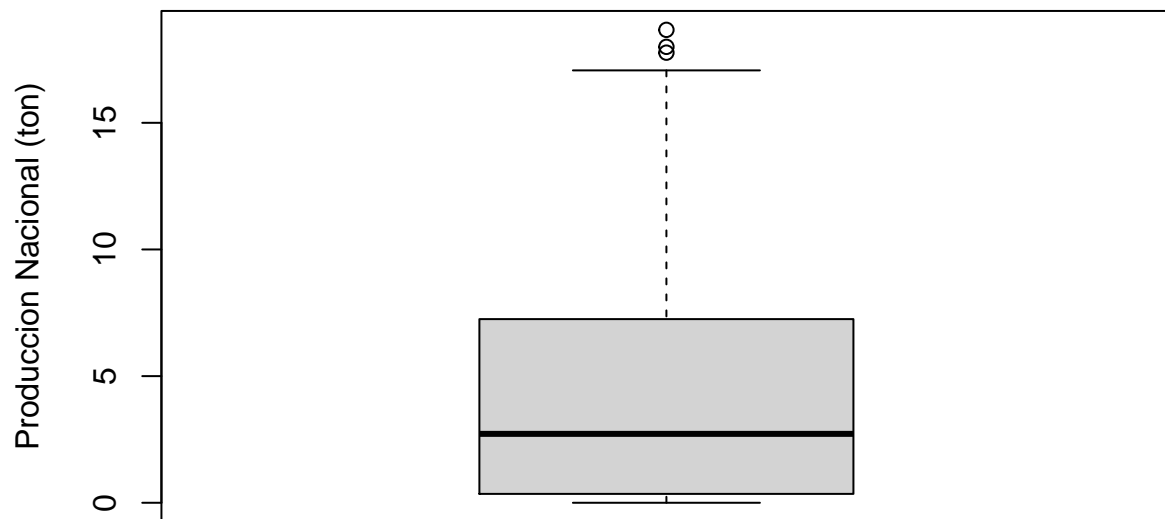


#LOS BOXPLOTS O DIAGRAMAS DE CAJA SE CONTRIBUYEN A PARTIR DE LOS PERCENTILES. #SE CONSTRUYE UN RECTANGULO USANDO ENTRE EL PRIMER Y EL TERCER CUARTIL. #LA ALTURA DEL RECTANGULO ES EL RANGO INTERCUARTIL. #LA MEDIANA ES UNA LINEA QUE DIVIDE EL RECTANGULO- #LOS VALORES MAS EXTREMOS QUE EL LARGO DE LOS BRAZOS SON CONSIDERADOS ATIPICOS #EL BOXPLOT NOS ENTREGA INFORMACION SOBRE LA SIMETRIA DE LA DISTRIBUCION DE LOS DATOS, SI LA MEDIANA NO ESTA EN EL CENTRO DEL RECTANGULO LA DISTRIBUCION NO ES SIMETRICRA. #SON UTILES PARA VER LA PRESENCIA DE VALORES ATIPICOS O OUTLERS

```
boxplot(cafe_df$Rendimiento..ha.ton.,ylab="Rendimiento (ha/ton)")
```



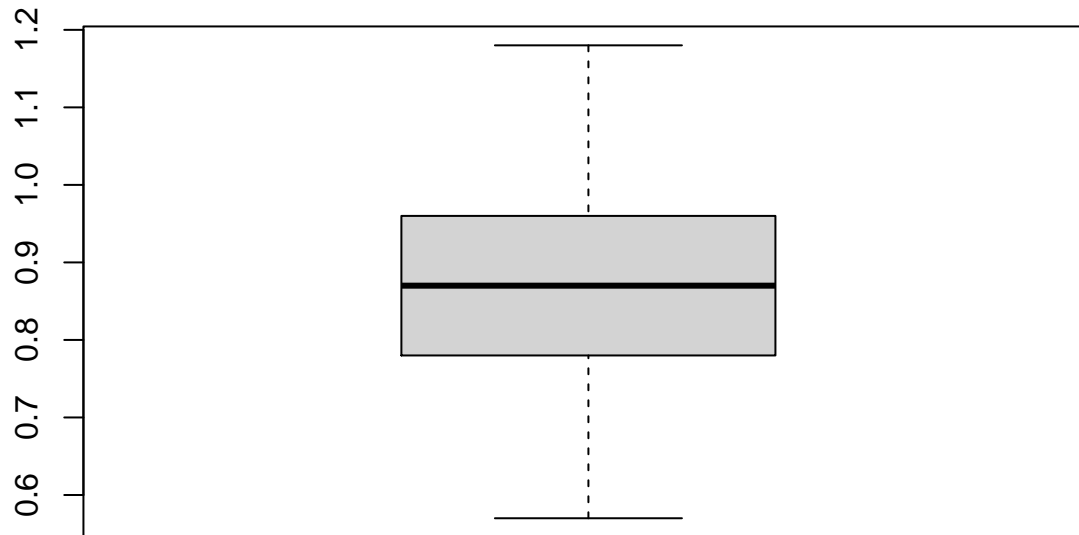
```
boxplot(caffe_df$Produccion.Nacional..ton.,ylab="Produccion Nacional (ton)")
```



```
#SELECCIONAR POR RANGOS
```

```
boxplot(x=caffe_df$Rendimiento..ha.ton.[1:10],main="Rendimiento (ha/ton)")
```

## Rendimiento (ha/ton)

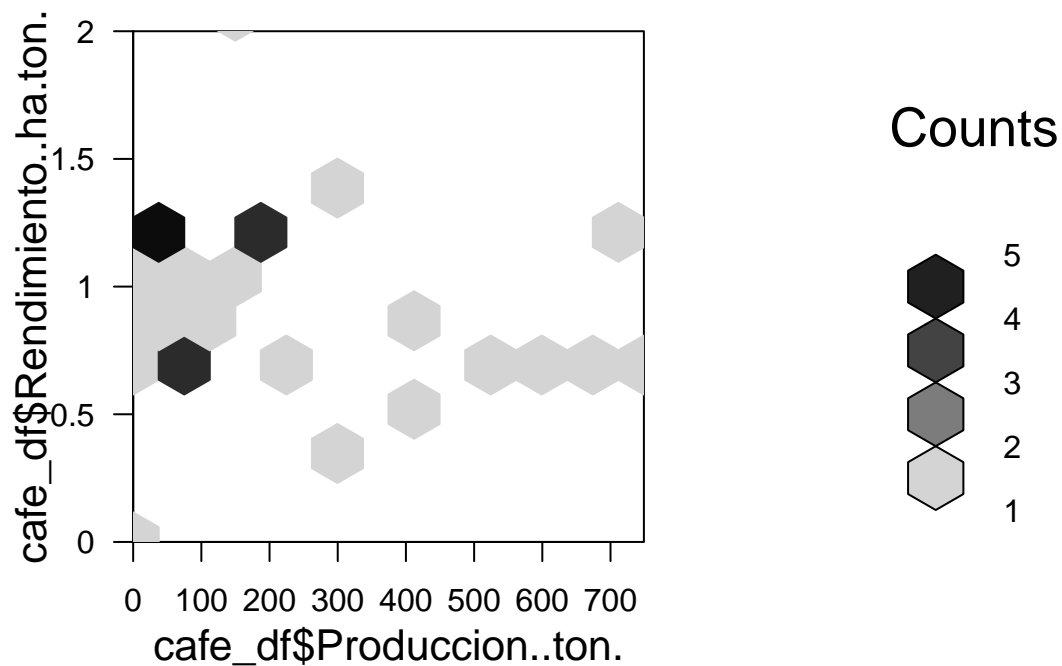


#CON LA FUNCION HEXBIN SE PUEDE MEJORAR LA PRESENTACION Y LECTURA DE LOS GRAFICOS. #SE CREA EL OBJETO BIN CON LA RELACION DE DOS VARIABLES

```
bin1<-hexbin(cafe_df$Produccion..ton.,cafe_df$Rendimiento..ha.ton., xbins=10)
```

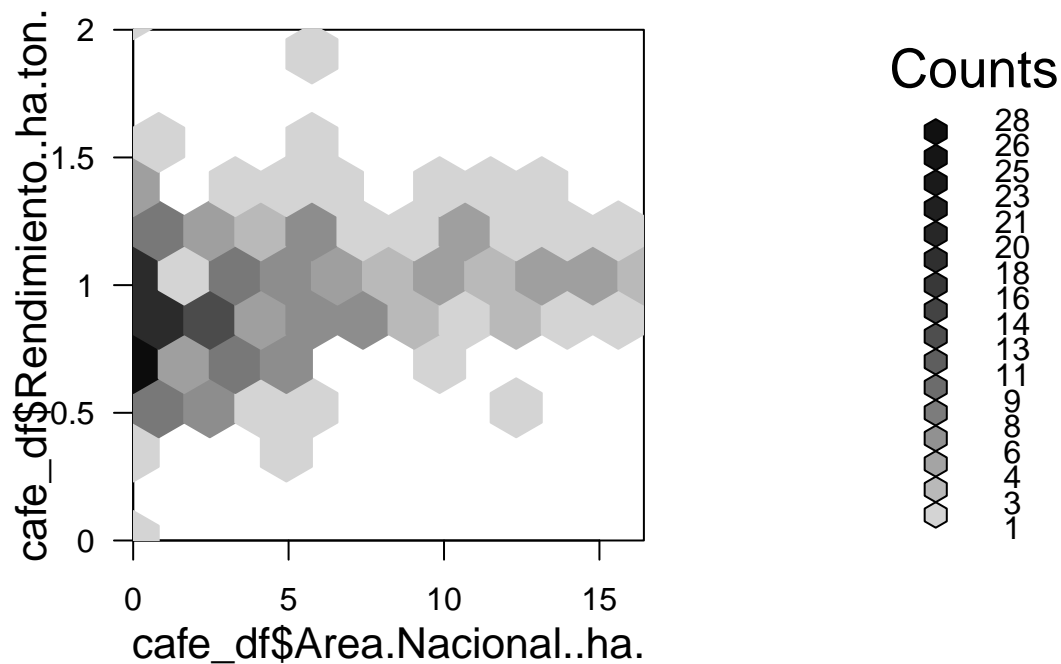
```
## Warning in xy.coords(x, y, xl, yl): NAs introduced by coercion
```

```
plot(bin1)
```



```
bin2<-hexbin(cafe_df$Area.Nacional..ha.,cafe_df$Rendimiento..ha.ton., xbins=10)
```

```
plot(bin2)
```



#HALLAMOS LA CORRELACION ENTRE LAS VARIABLES POR CORRELACION DE PEARSON-  
#INDICAMOS LAS VARIABLES CON VALORES DESCARTANDO LOS VALORES NULOS QUE EXISTAN

```
cor(cafe_df$Area.Nacional..ha.,cafe_df$Rendimiento..ha.ton., use="complete.obs")
```

```
## [1] 0.2806767
```

#PARA CALCULAR LA MODA SELECCIONANDO LAS VARIABLES EN EL DATAFRAME “cafe\_df”

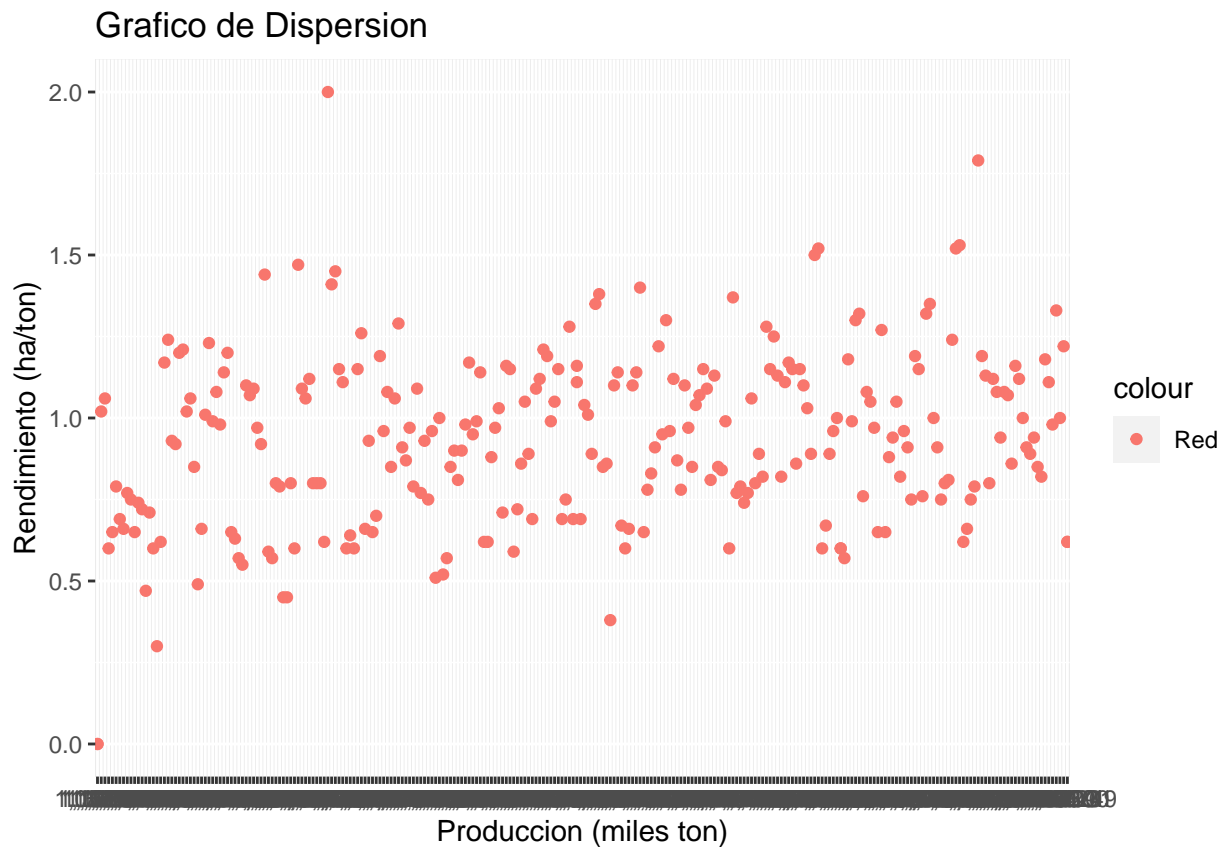
```
moda=function(var){  
  frec.var<-table(var)  
  valor=which(frec.var==max(frec.var)) #Elemento con el valor  
  names(valor)}
```

```
which.max(cafe_df$Rendimiento..ha.ton.)
```

```
## [1] 121
```

#SEGUIMOS ESTUDIANDO EL COMPORTAMIENTO O DISTRIBUCION DE LOS DATOS O LA INFORMACION CON GRAFICOS DE DISPERSION. #GRAFICO DE DISPERSION DEL COMPORTAMIENTO DE LA PRODUCCION VS RENDIMIENTO, PERMITE ENTENDER O DESCUBRIR PATRONES EN LOS DATOS

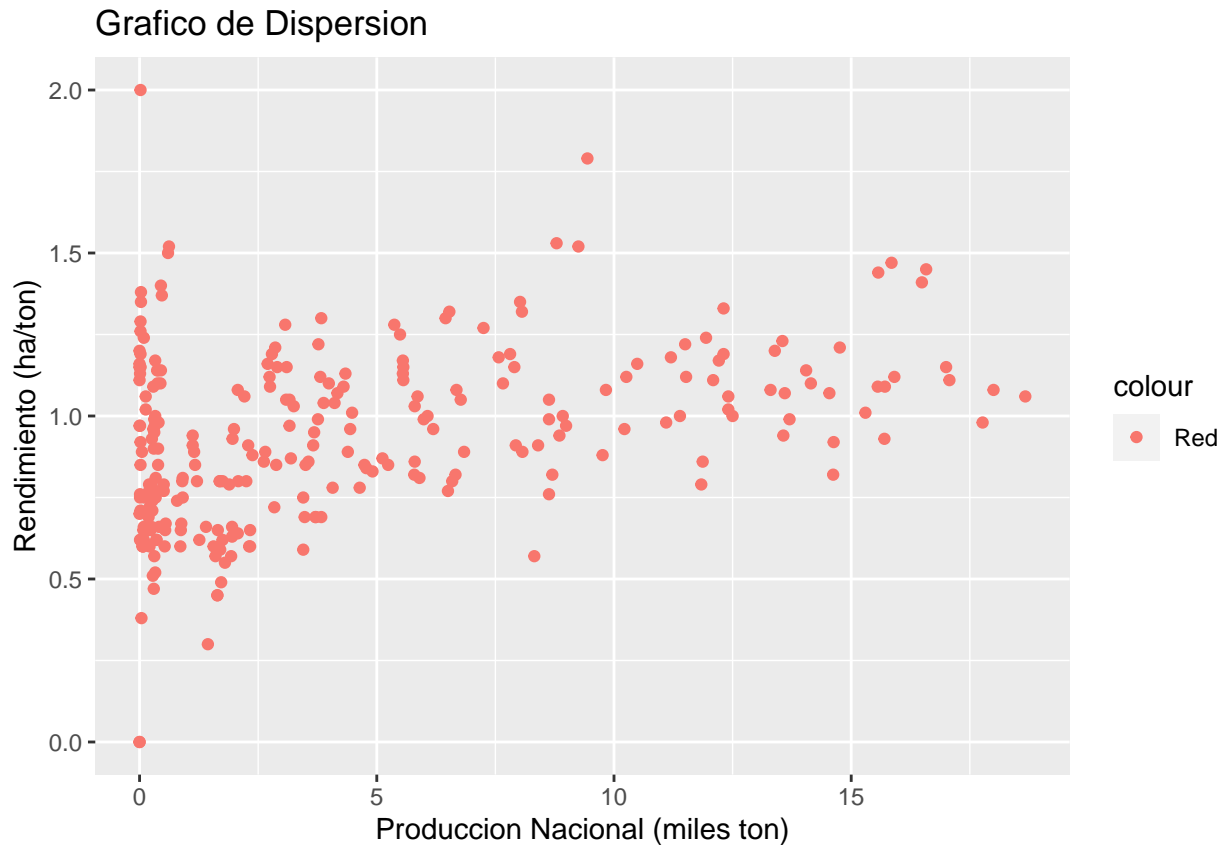
```
qplot(Produccion..ton., Rendimiento..ha.ton., data = cafe_df,  
      main = "Grafico de Dispersion",  
      col = "Red",  
      xlab = "Produccion (miles ton)",  
      ylab = "Rendimiento (ha/ton)")
```



#GRAFICO DE DISPERSION DEL COMPORTAMIENTO DE LA PRODUCCION NACIONAL VS RENDIMIENTO.

```
qplot(Produccion.Nacional..ton., Rendimiento..ha.ton., data = cafe_df,  
      main = "Grafico de Dispersion",  
      col = "Red",  
      xlab = "Produccion Nacional (miles ton)",  
      ylab = "Rendimiento (ha/ton)")
```





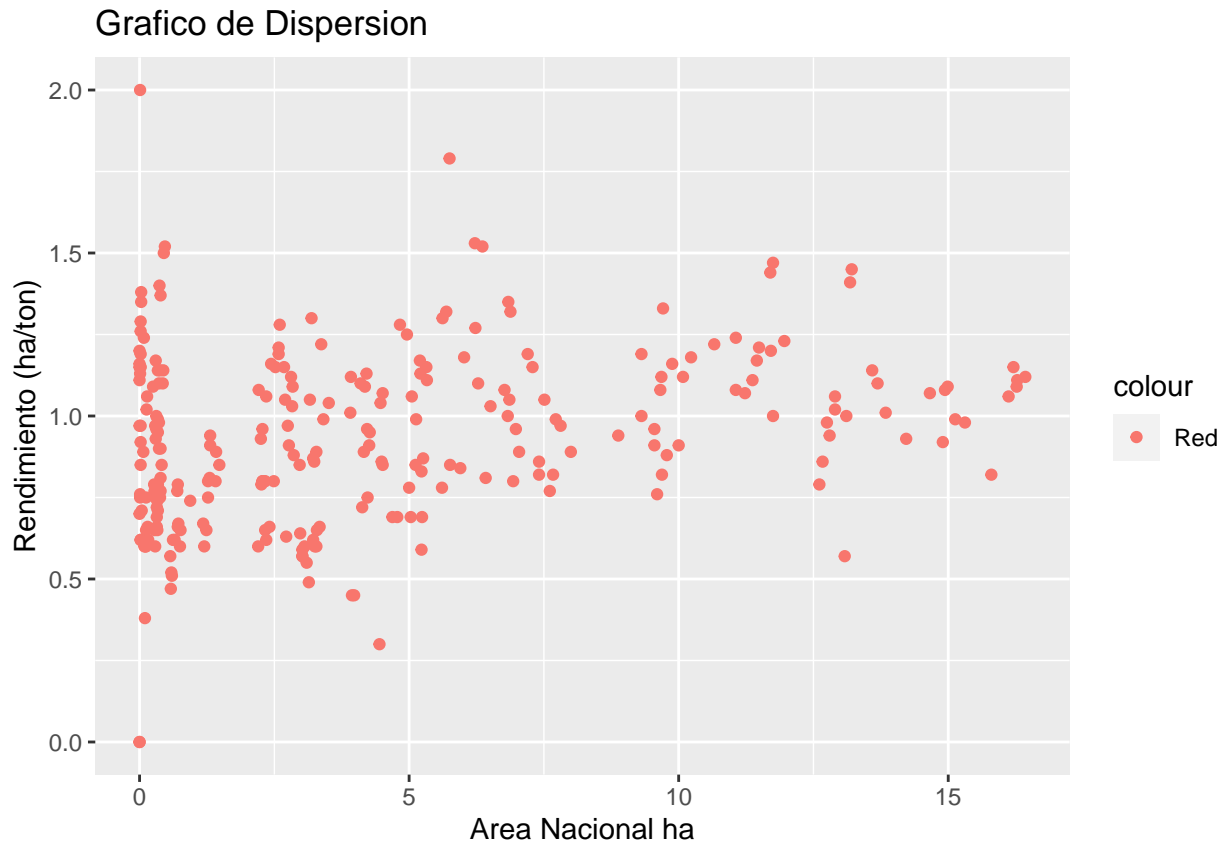
#COEFICIENTE DE CORRELACION DE PEARSON ENTRE LA PRODUCCION NACIONAL Y EL RENDIMIENTO. #HALLAS LA CORRELACION ENTRE LAS VARIABLES POR CORRELACION DE PEARSON #INDICAMOS QUE UTILIZA LAS VARIABLES CON VALORES DESCARTANDO LOS VALORES NULOS QUE EXISTAN

```
cor(cafe_df$Produccion.Nacional..ton., cafe_df$Rendimiento..ha.ton., use="complete.obs")
```

```
## [1] 0.3855697
```

#GRAFICO DE DISPERSION DEL COMPORTAMIENTO ENTRE EL AREA NACIONAL VS EL RENDIMIENTO. PERMITE ENTENDER O DESCUBRIR PATRONES EN LOS DATOS

```
qplot(Area.Nacional..ha., Rendimiento..ha.ton., data = cafe_df,
      main = "Grafico de Dispersion",
      col = "Red",
      xlab = "Area Nacional ha",
      ylab = "Rendimiento (ha/ton)")
```

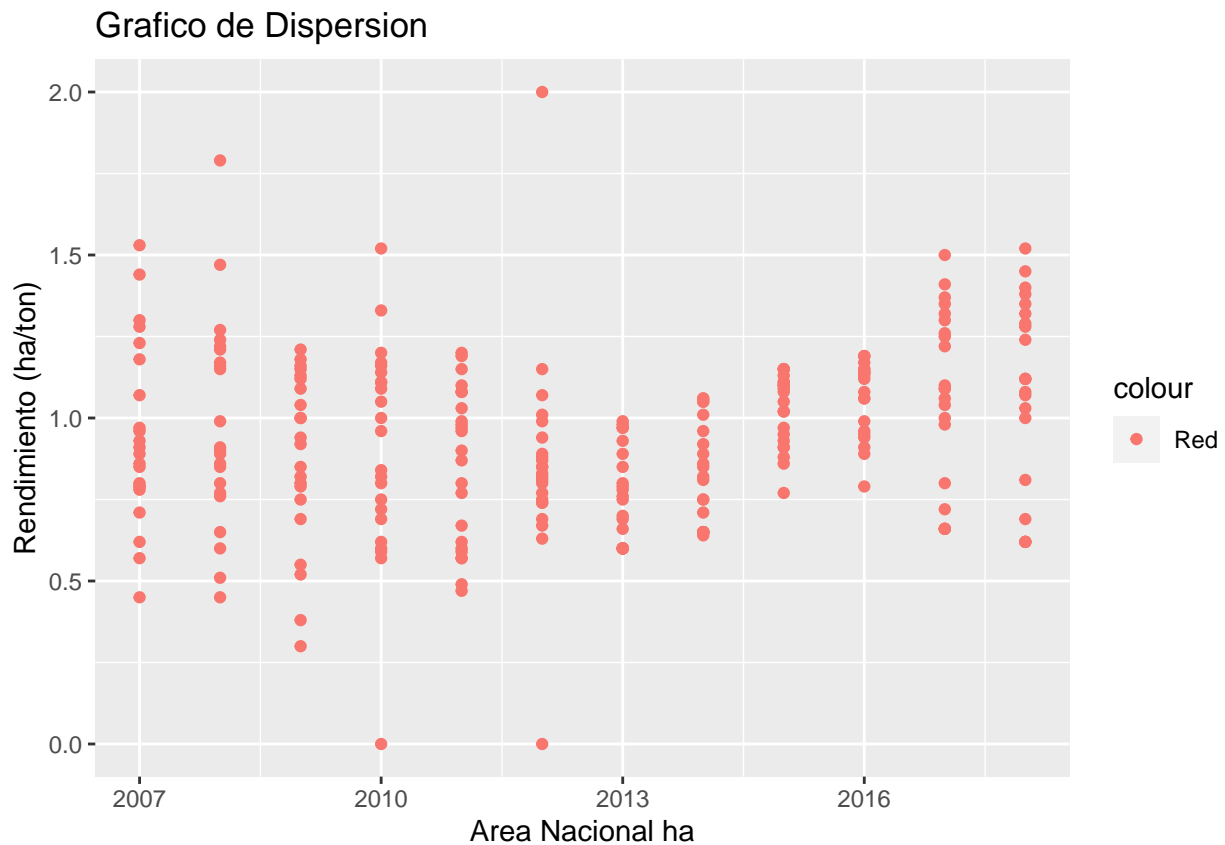


#COEFICIENTE DE CORRELACION DE PEARSON #COEFICIENTE DE CORRELACIÓN ENTRE EL AREA NA-CIONAL Y EL RENDIMIENTO #HALLAR LA CORRELACION ENTRE LAS VARIABLES POR METODO CORRELACION DE PEARSON #INDICAMOS QUE UTILIZA LAS VARIABLES CON VALORES DESCARTANDO LOS VALORES NULOS QUE EXISTAN

```
cor(cafe_df$Area.Nacional..ha.,cafe_df$Rendimiento..ha.ton., use="complete.obs")
```

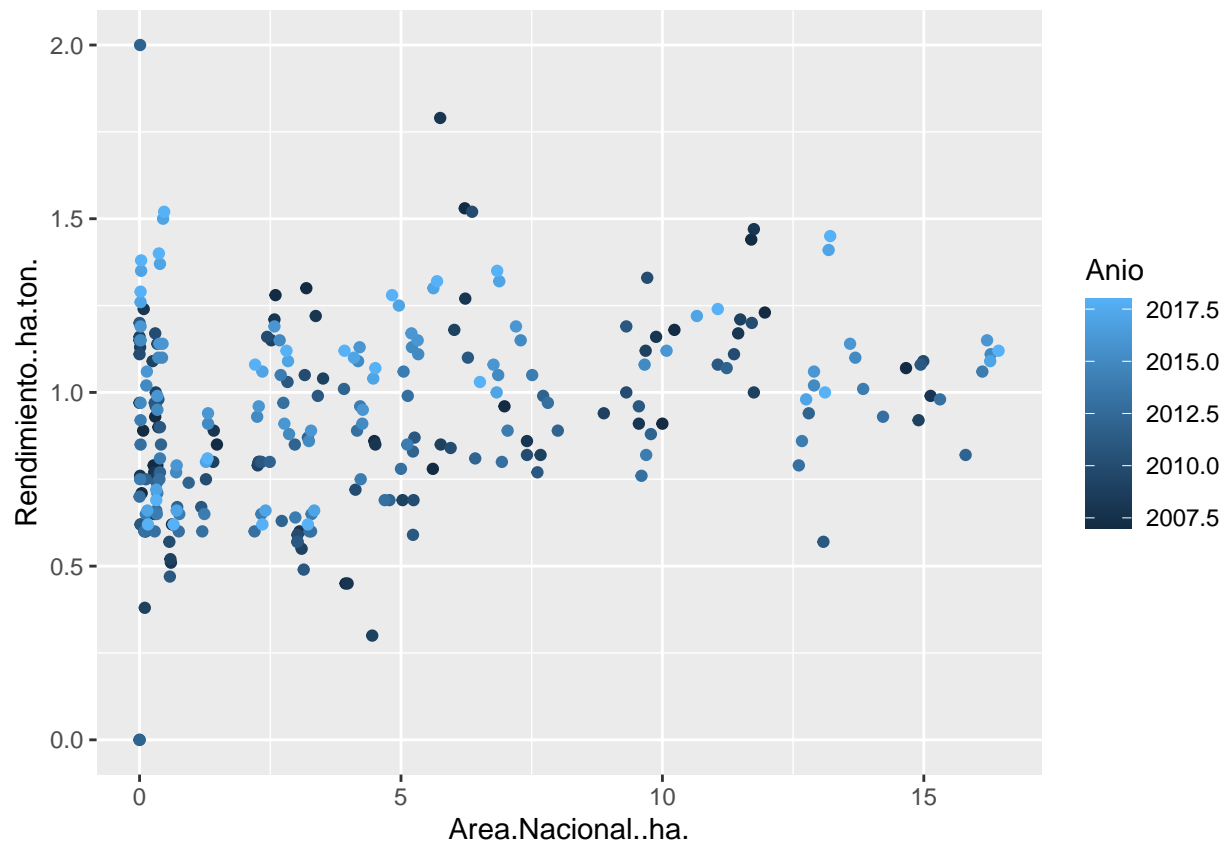
```
## [1] 0.2806767
```

```
qplot(Anio, Rendimiento..ha.ton., data = cafe_df,
      main = "Grafico de Dispersion",
      col = "Red",
      xlab = "Area Nacional ha",
      ylab = "Rendimiento (ha/ton)")
```



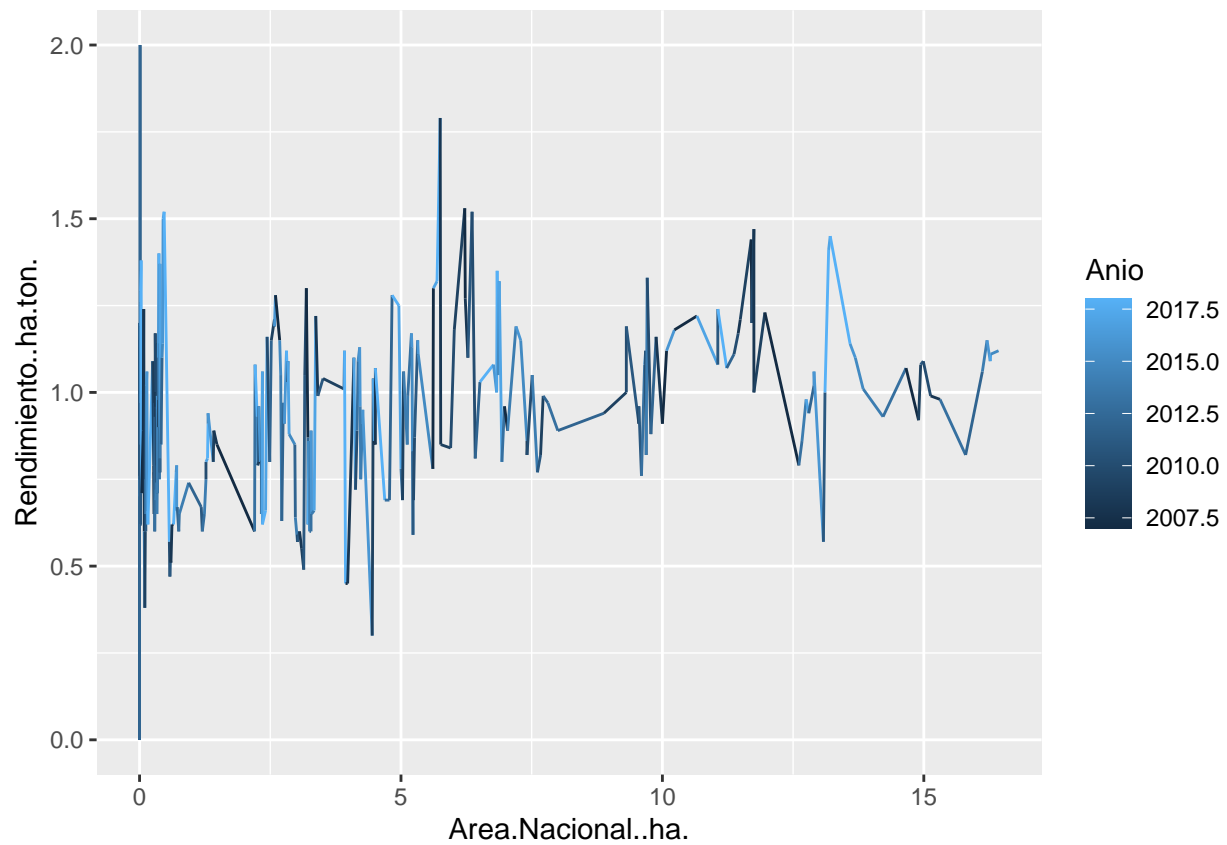
#ESPECIFICAR EL DATAFRAME ES EL PRIMER ARGUMENTO EN LA FUNCIÓN GGLOT #  
 DENTRO DE AES() ESCRIBIMOS LAS VARIABLES (X,Y) QUE QUEREMOS GRAFICAR # PERMITE  
 ENTENDER O DESCUBRIR PATRONES EN LOS DATOS # SE PUEDE UTILIZAR EL PARAMETRO  
 COLOR PARA REPRESENTAR UNA TERCERA VARIABLE # GRÁFICO DE DISPERSION: AREA  
 NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_point(aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton., color = Anio))
```



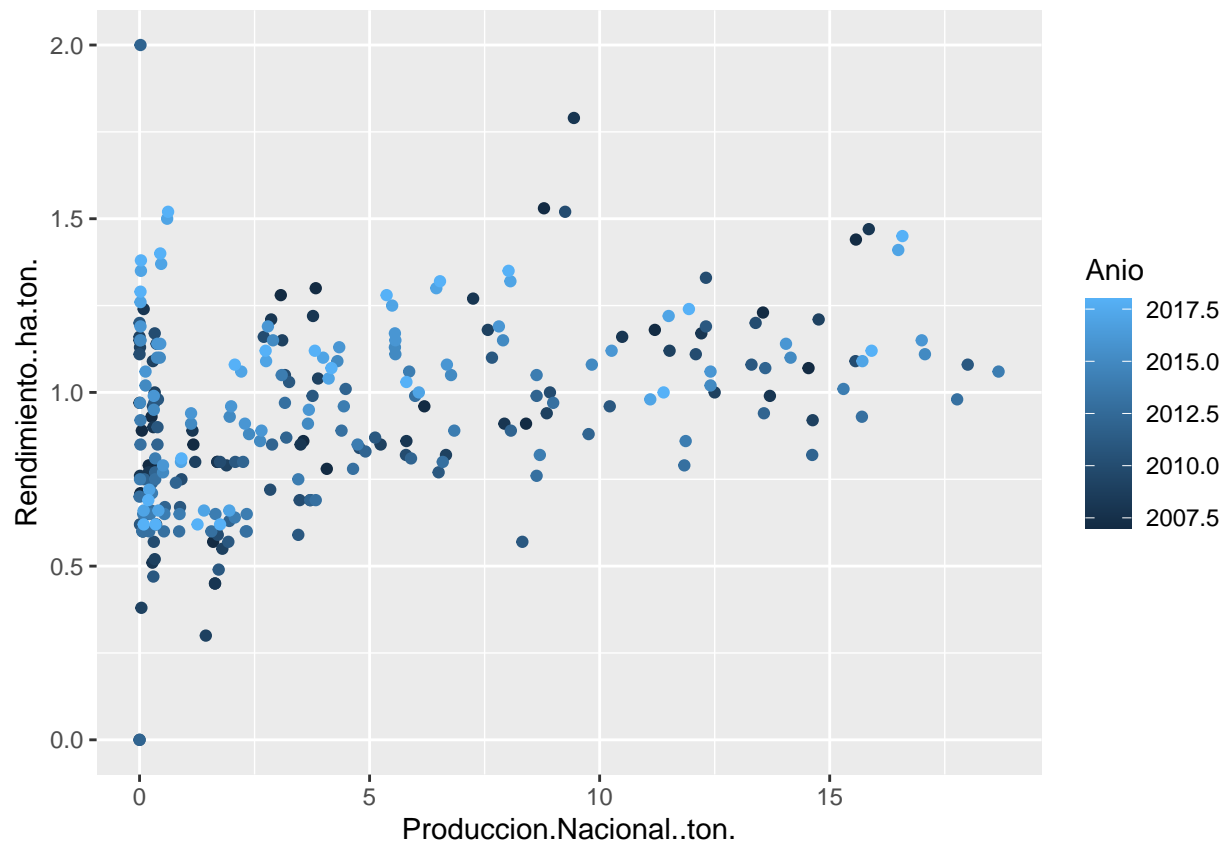
#GRAFICO DE LINEAS: AREA NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_line(aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton., color = Anio))
```



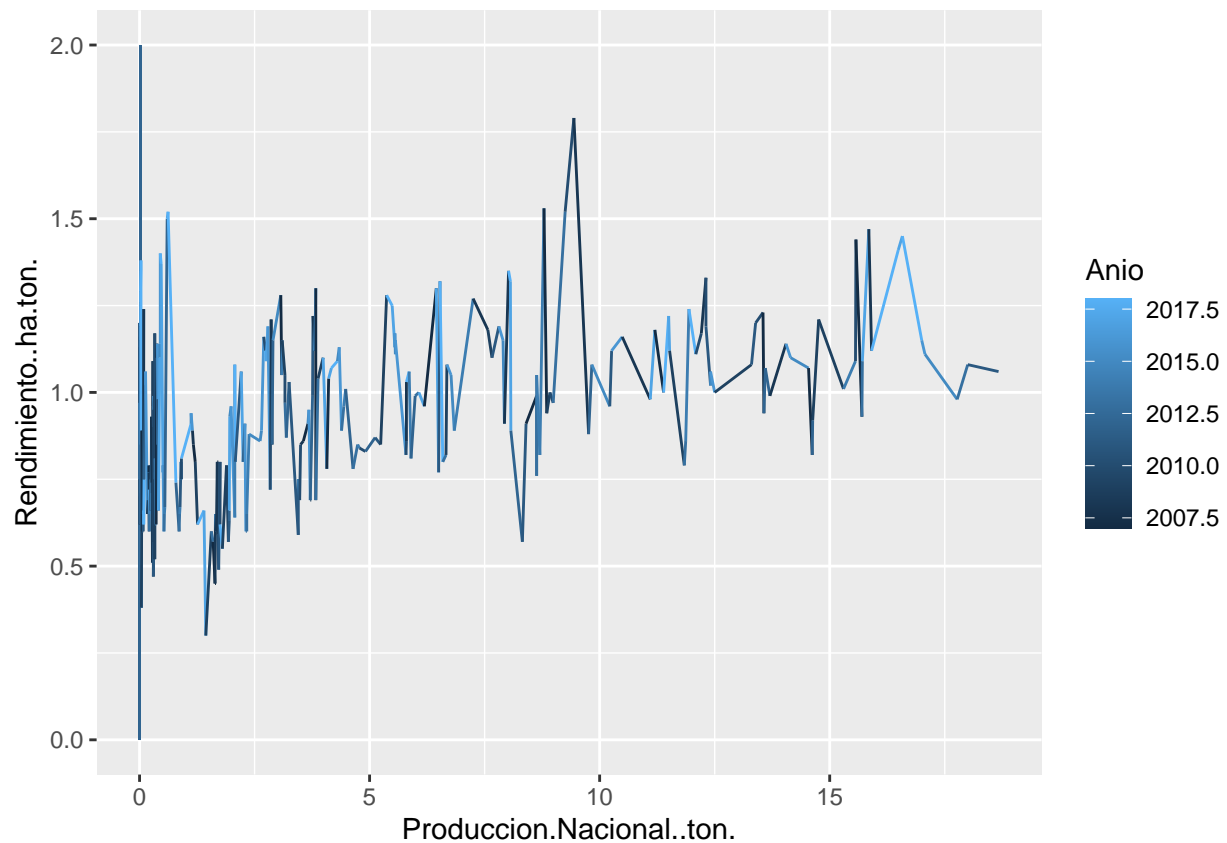
#GRAFICO DE DISPERSION: PRODUCCION NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_point(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio))
```



#GRAFICO DE LINEAS: PRODUCCION NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_line(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio))
```

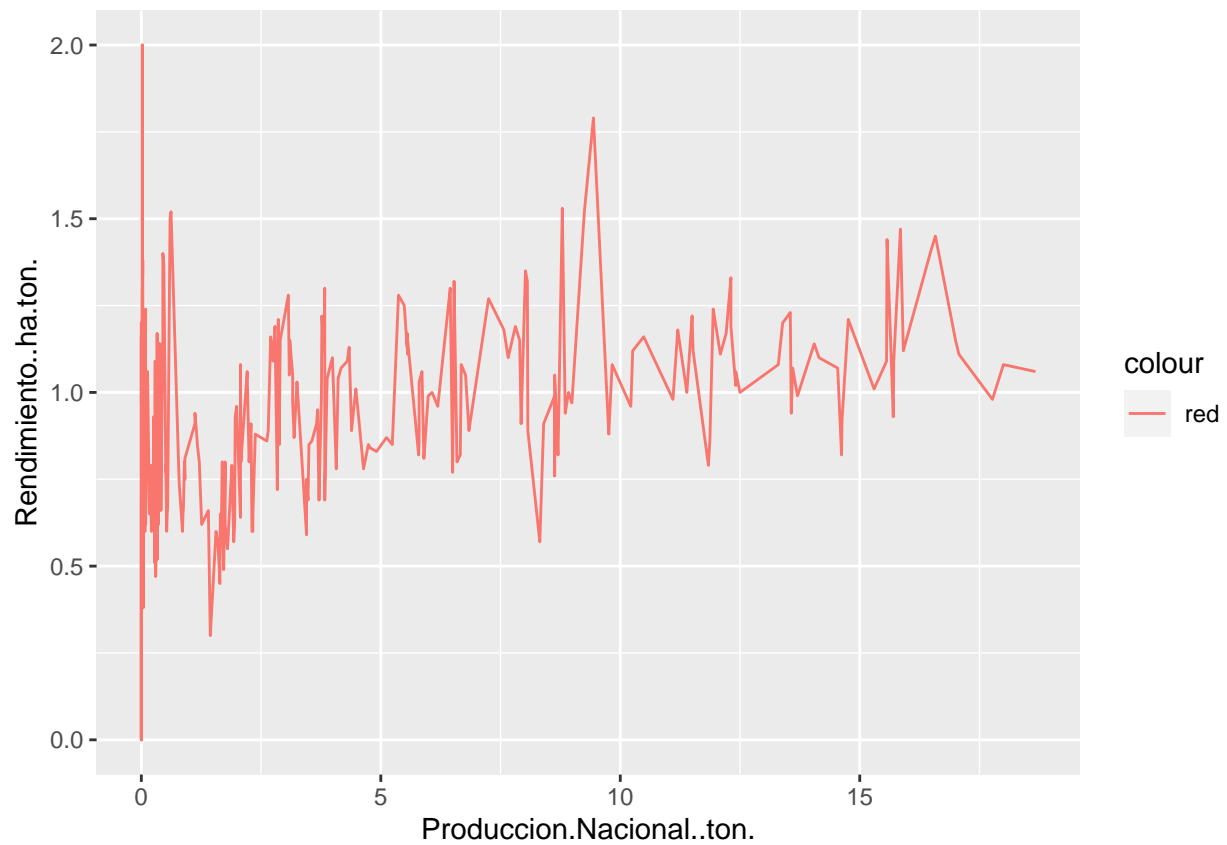


#OTRA FORMA DE GRAFICO DE LINEAS SIN REPRESENTAR LOS AÑOS

```
data(cafe_df)
```

```
## Warning in data(cafe_df): data set 'cafe_df' not found
```

```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = "red")) +  
  geom_line()
```

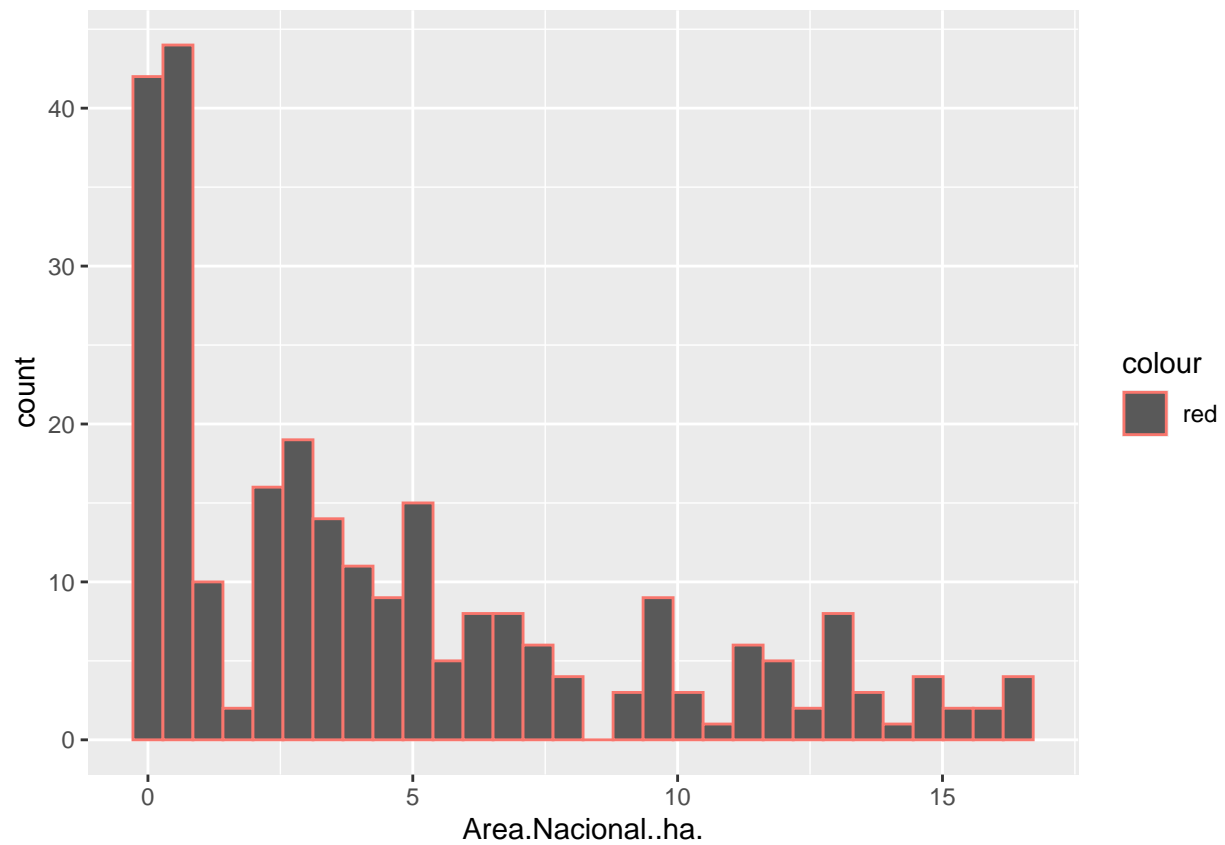


#HISTOGRAMA: AREA NACIONAL

```
ggplot(cafe_df) +  
  geom_histogram(aes(x = Area.Nacional..ha., color="red"))
```

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

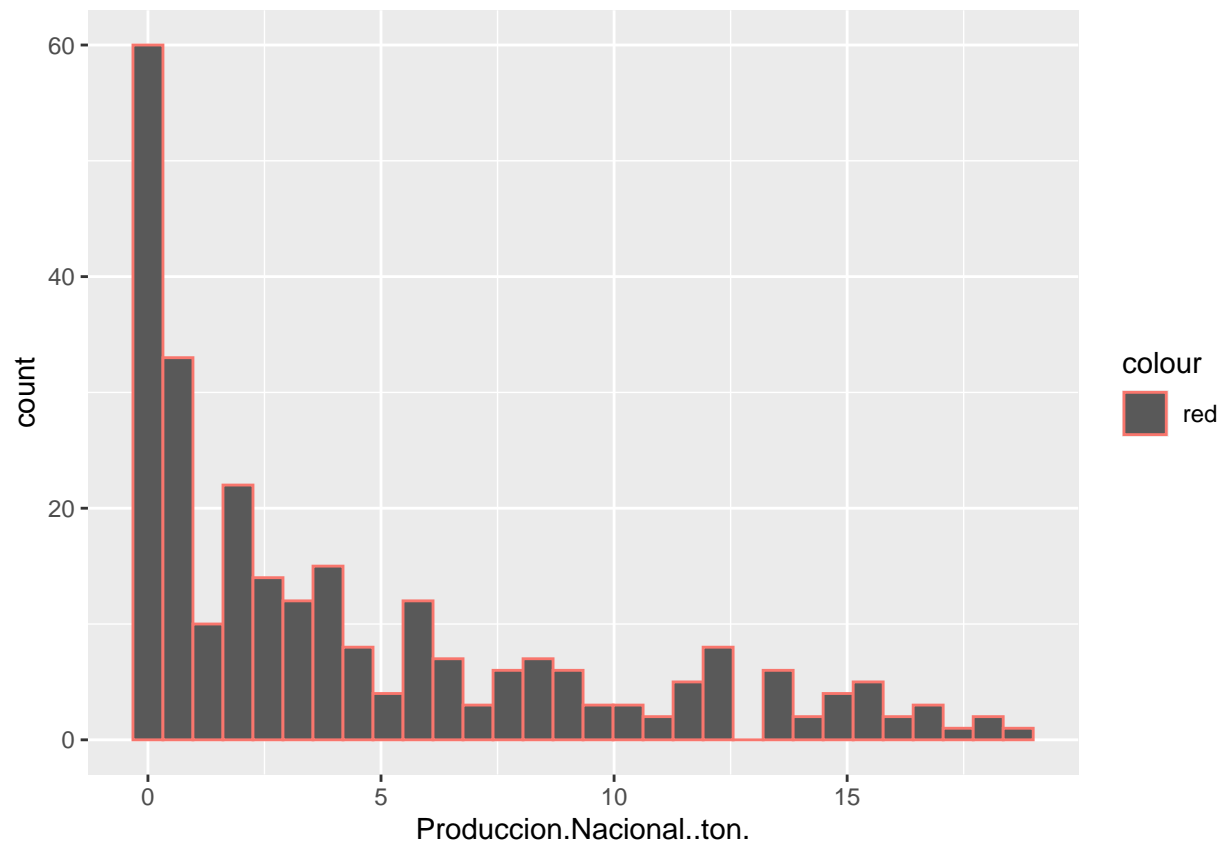




#HISTOGRAMA PRODUCCION NACIONAL

```
ggplot(cafe_df) +
  geom_histogram(aes(x = Produccion.Nacional..ton., color="red"))
```

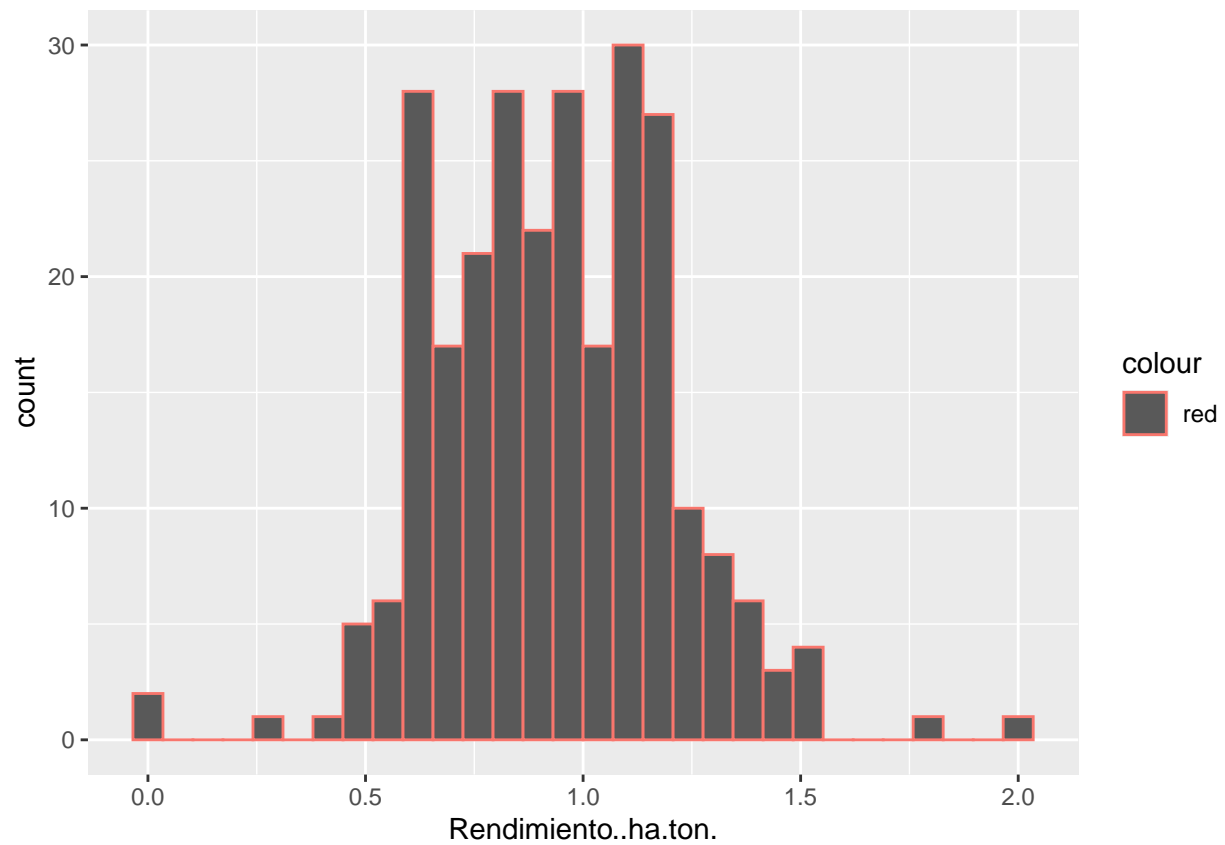
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



#HISTOGRAMA RENDIMIENTO

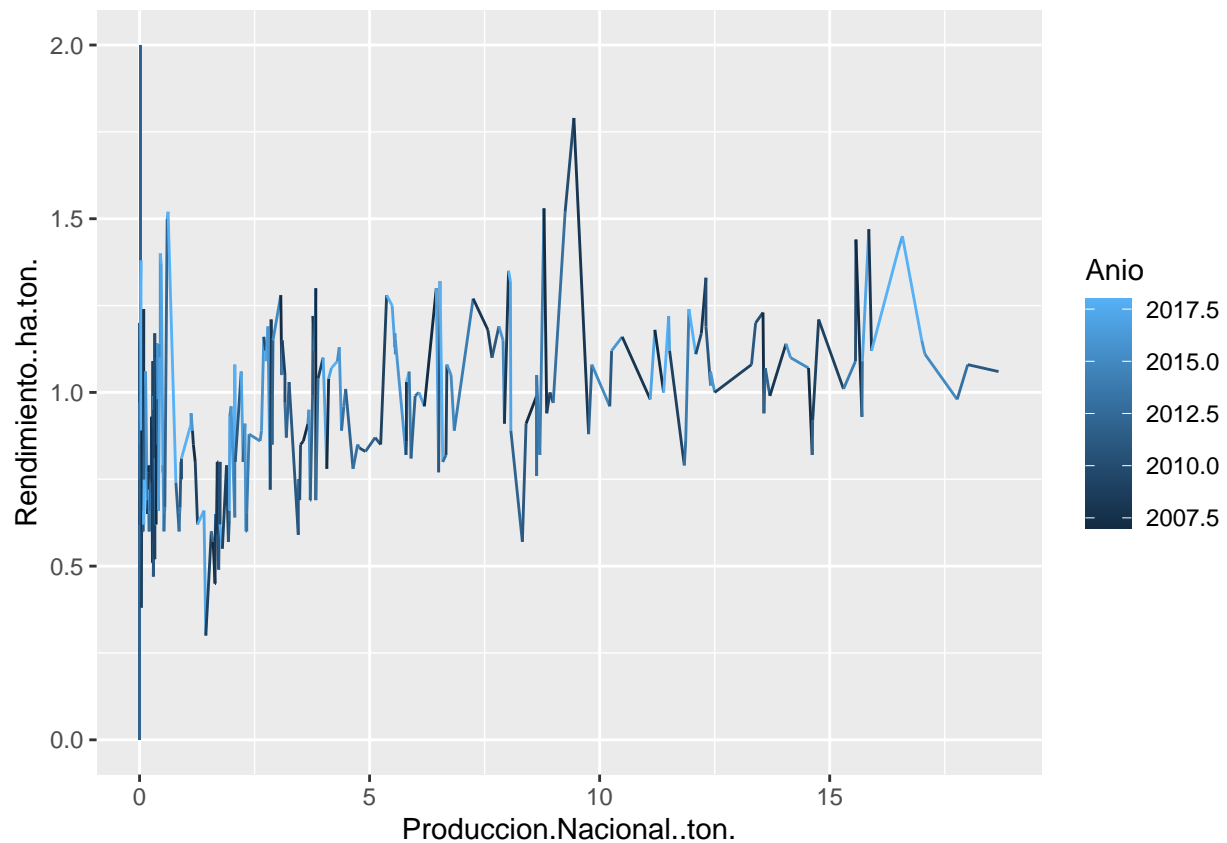
```
ggplot(cape_df) +  
  geom_histogram(aes(x = Rendimiento..ha.ton., color="red"))
```

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



#GRAFICO DE LINEAS: PRODUCCION NACIONAL VS RENDIMIENTO

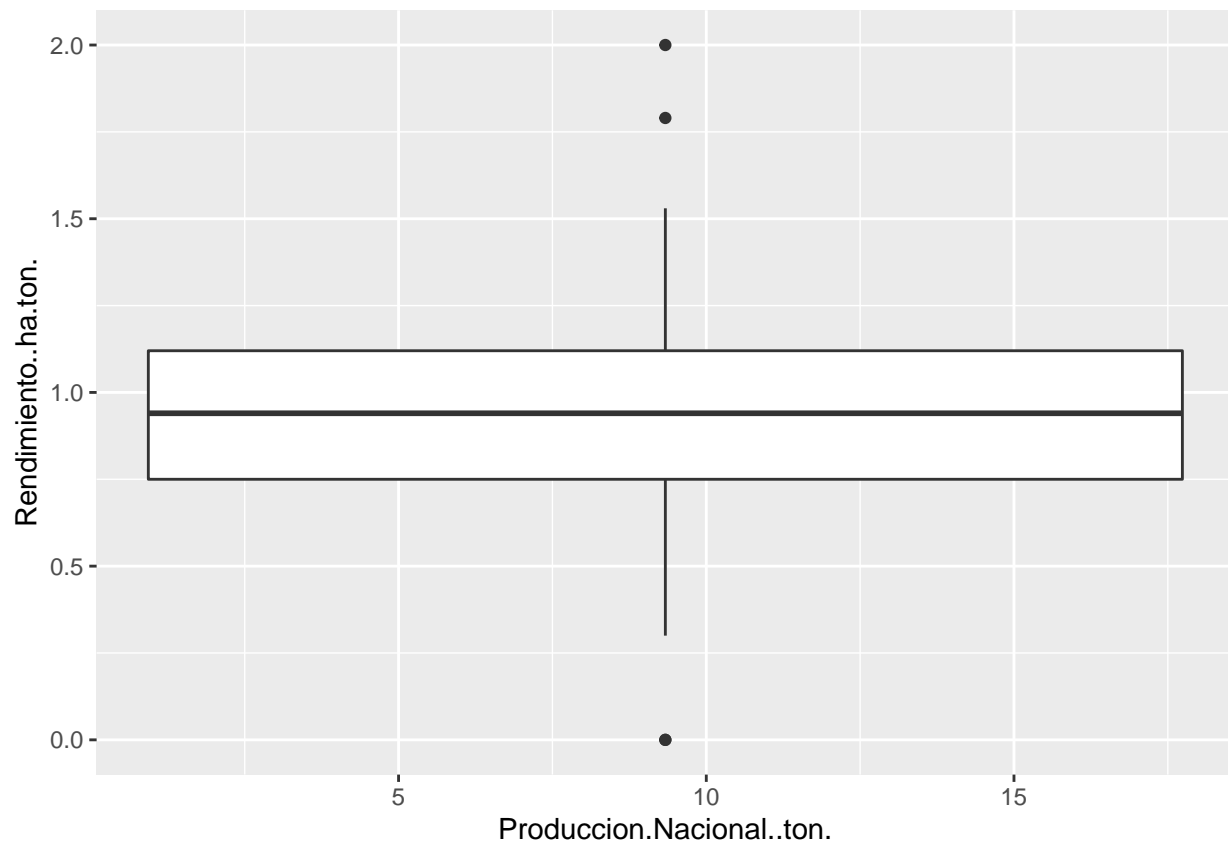
```
ggplot(cape_df) +
  geom_line(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio))
```



#GRAFICO DE CAJAS: PRODUCCION NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_boxplot(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Año))
```

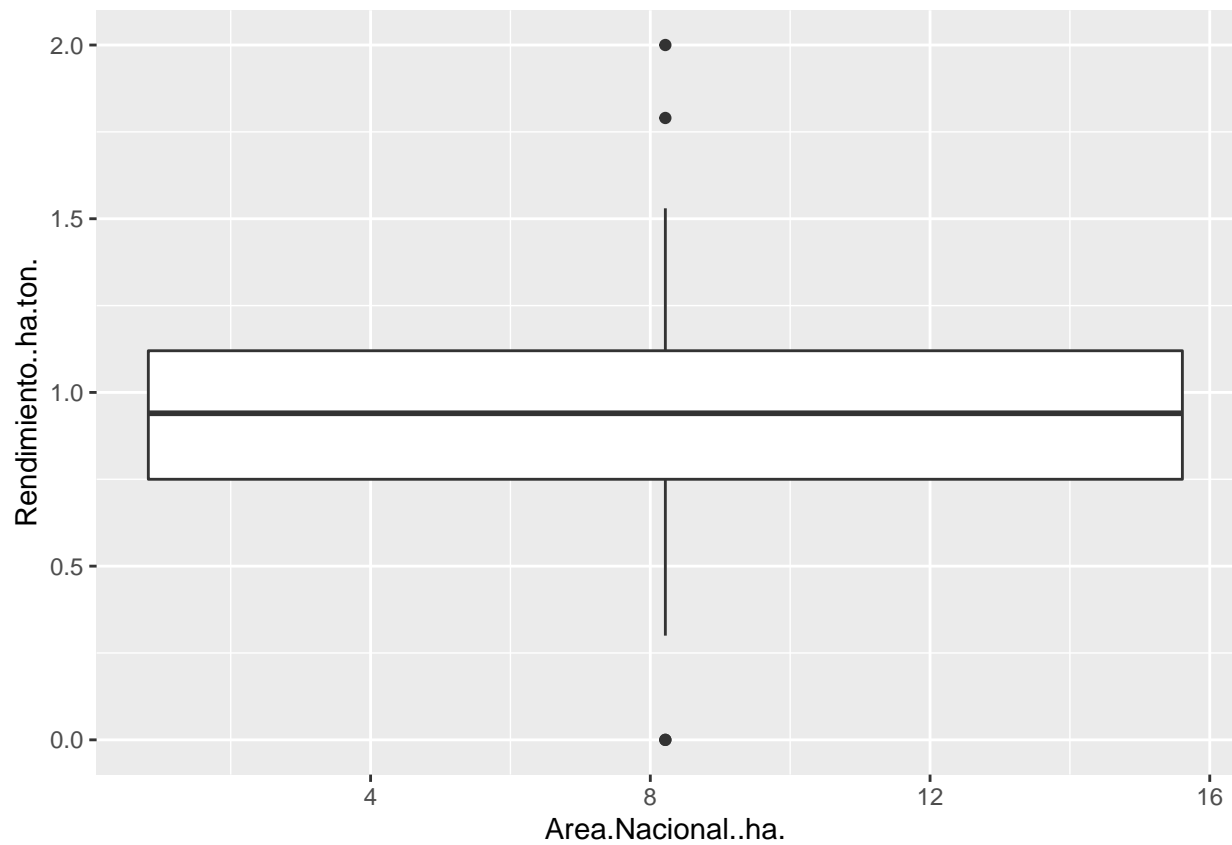
## Warning: Continuous x aesthetic -- did you forget aes(group=...)?



#GRAFICO DE CAJAS: AREA NACIONAS VS RENDIMIENTO

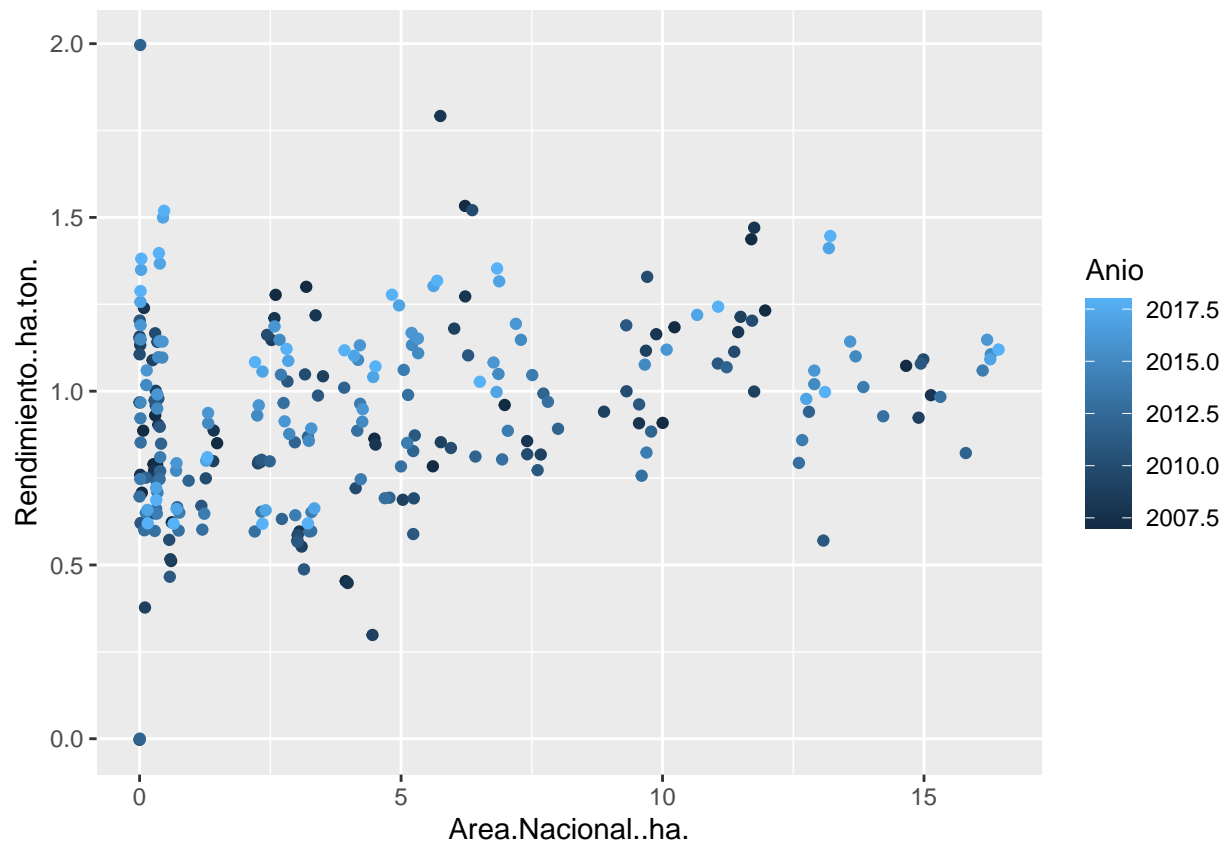
```
ggplot(cafe_df) +
  geom_boxplot(aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton., color = Anio))
```

## Warning: Continuous x aesthetic -- did you forget aes(group=...)?



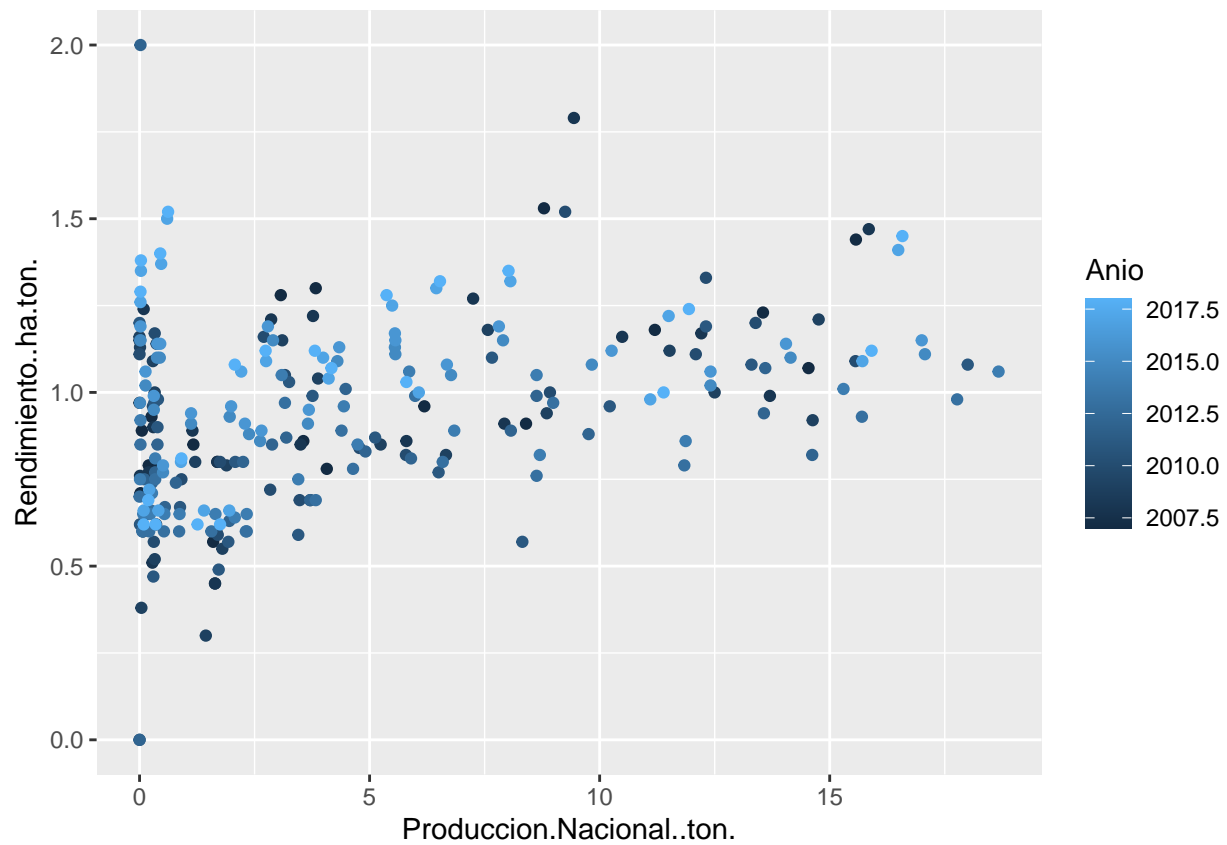
#GRAFICO DE DISPERSION JITTER: AREA NACIONAL VS RENDIMIENTO

```
ggplot(caffe_df) +  
  geom_jitter(aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton., color = Anio))
```



#GRAFICO DE DISPERSION PRODUCCION NACIONAL VS RENDIMIENTO

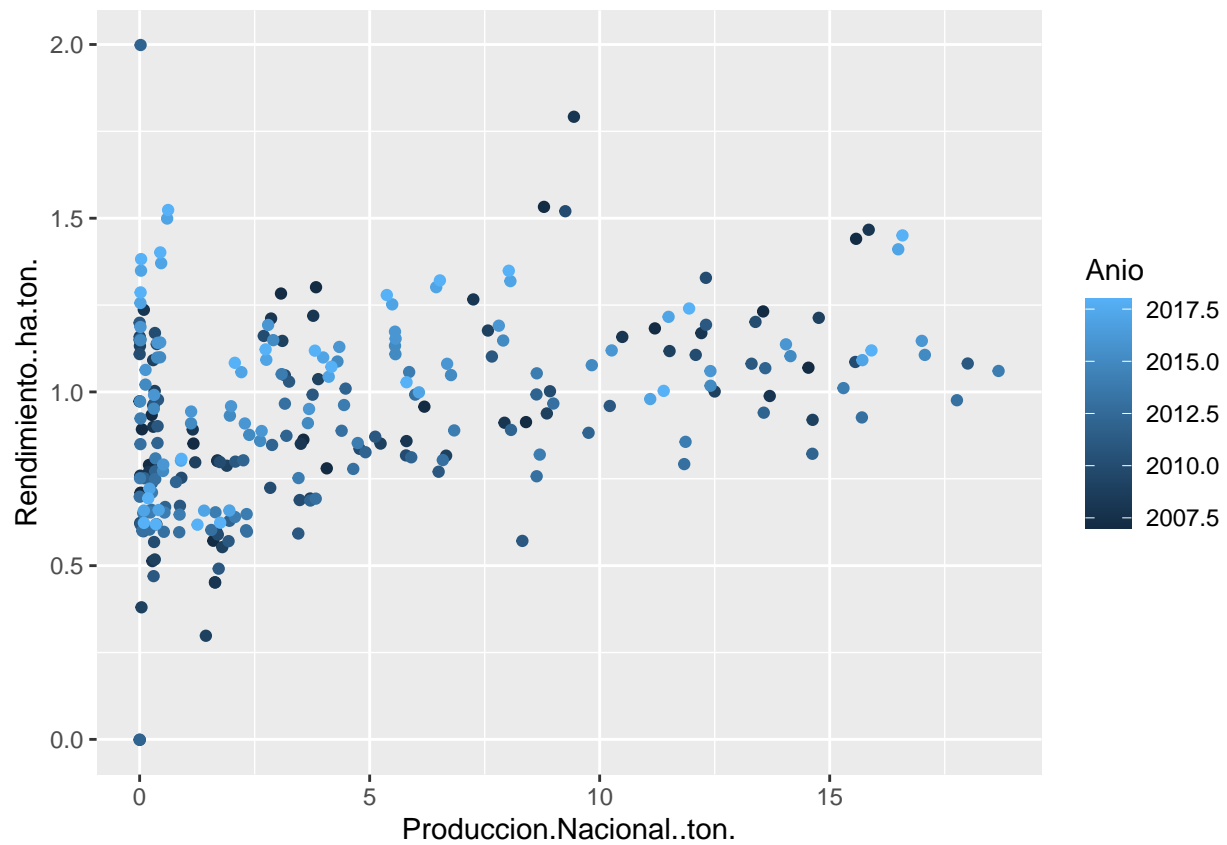
```
ggplot(cafe_df) +  
  geom_point(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio))
```



#GRAFICO DE DISPERSION JITTER: PRODUCCION NACIONAL VS RENDIMIENTO

```
ggplot(cafe_df) +  
  geom_jitter(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Año))
```



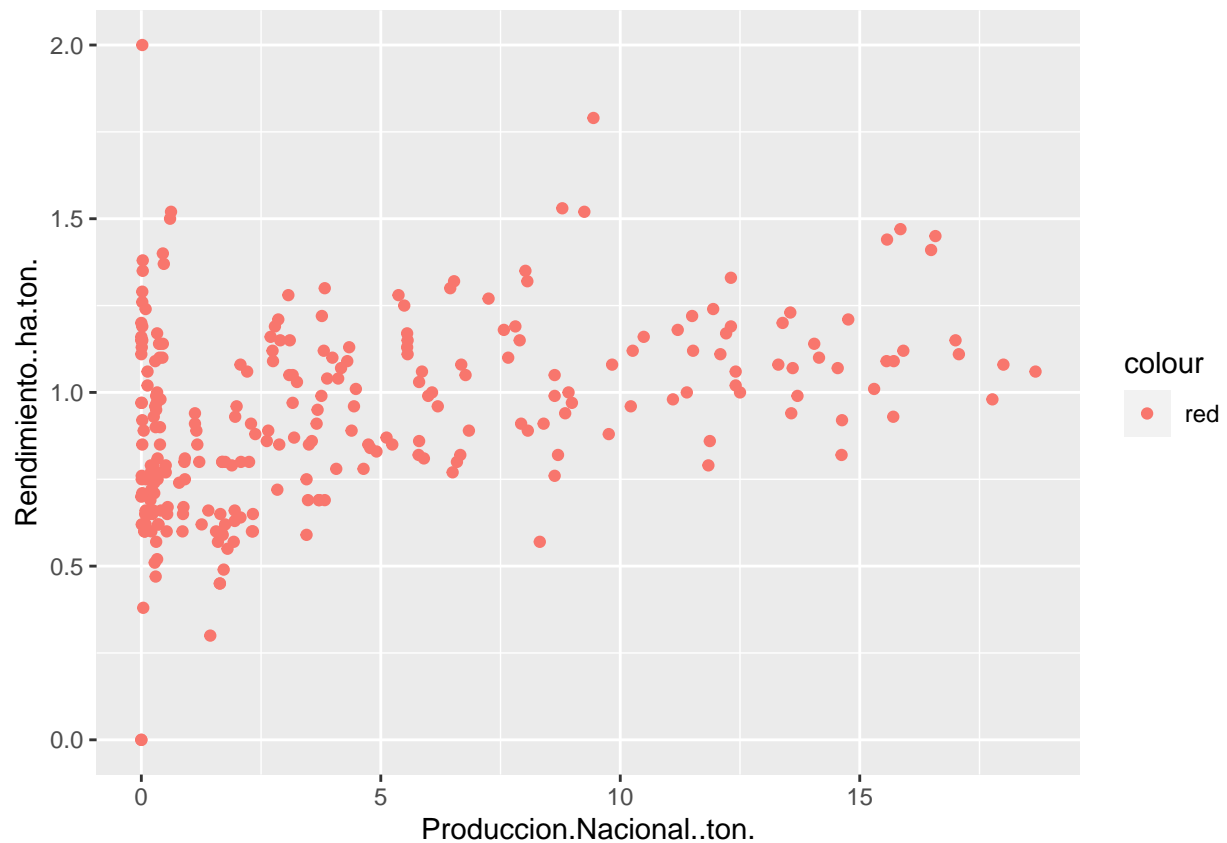


#OTRA FORMA CON GRAFICO DE DISPERSION SIN REPRESENTAR LOS AÑOS

```
data(cafe_df)
```

```
## Warning in data(cafe_df): data set 'cafe_df' not found
```

```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = "red")) +  
  geom_point()
```

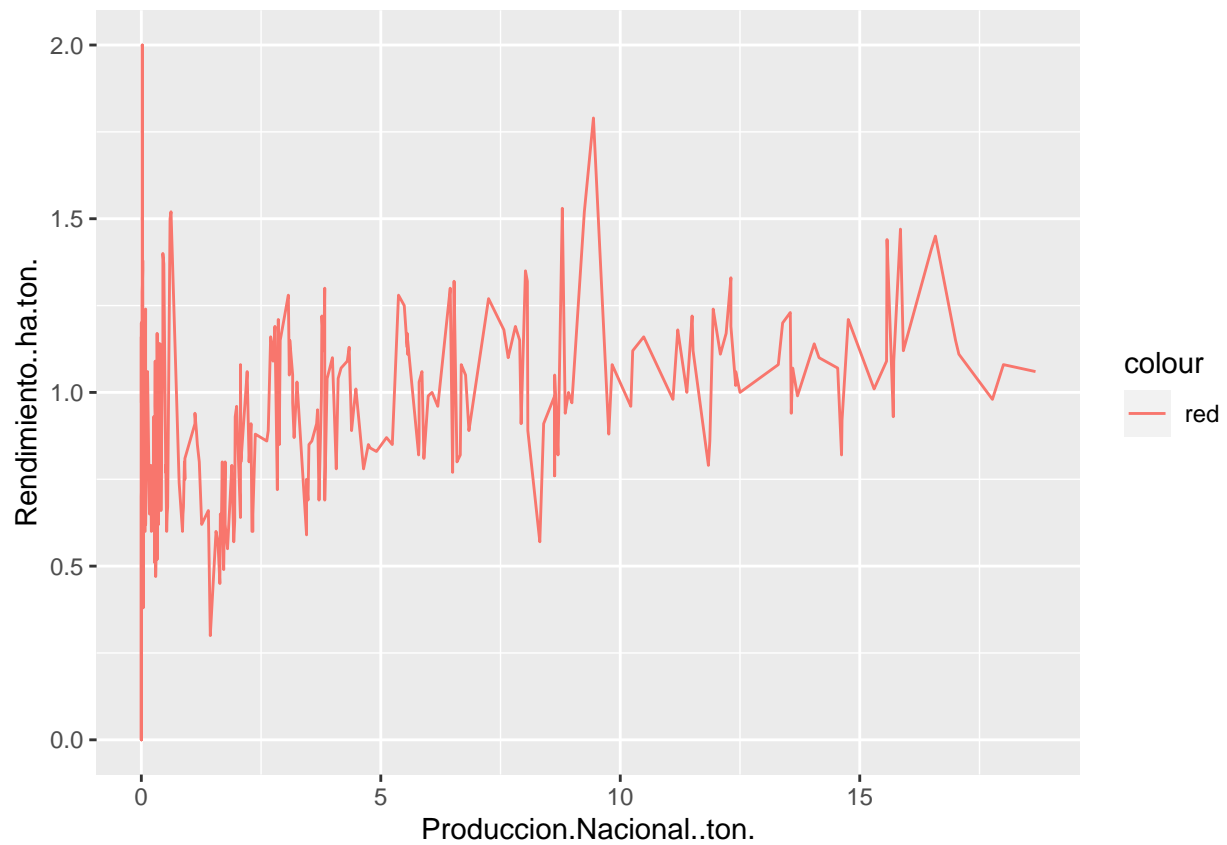


#OTRA FORMA CON GRAFICO DE LINEAS SIN REPRESENTAR LOS AÑOS

```
data(caffe_df)
```

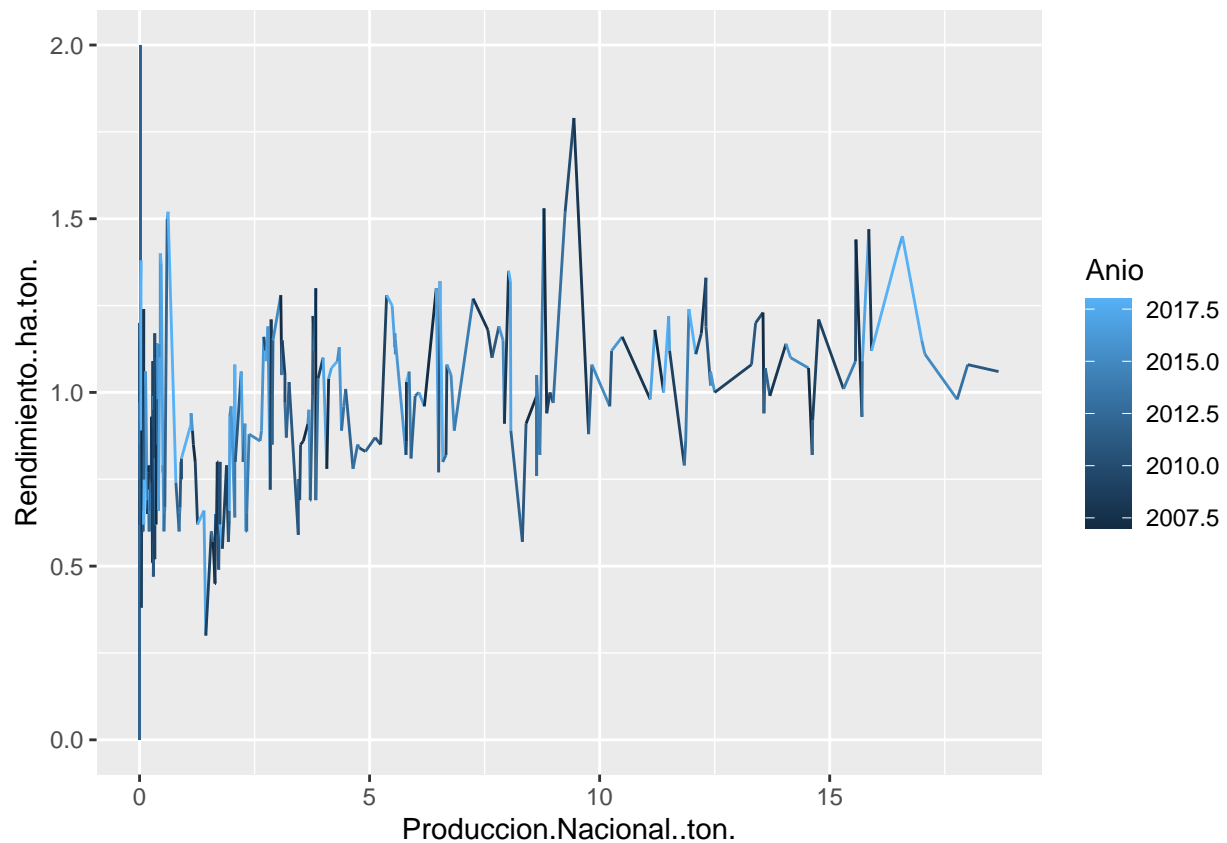
```
## Warning in data(caffe_df): data set 'caffe_df' not found
```

```
ggplot(caffe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = "red")) +  
  geom_line()
```



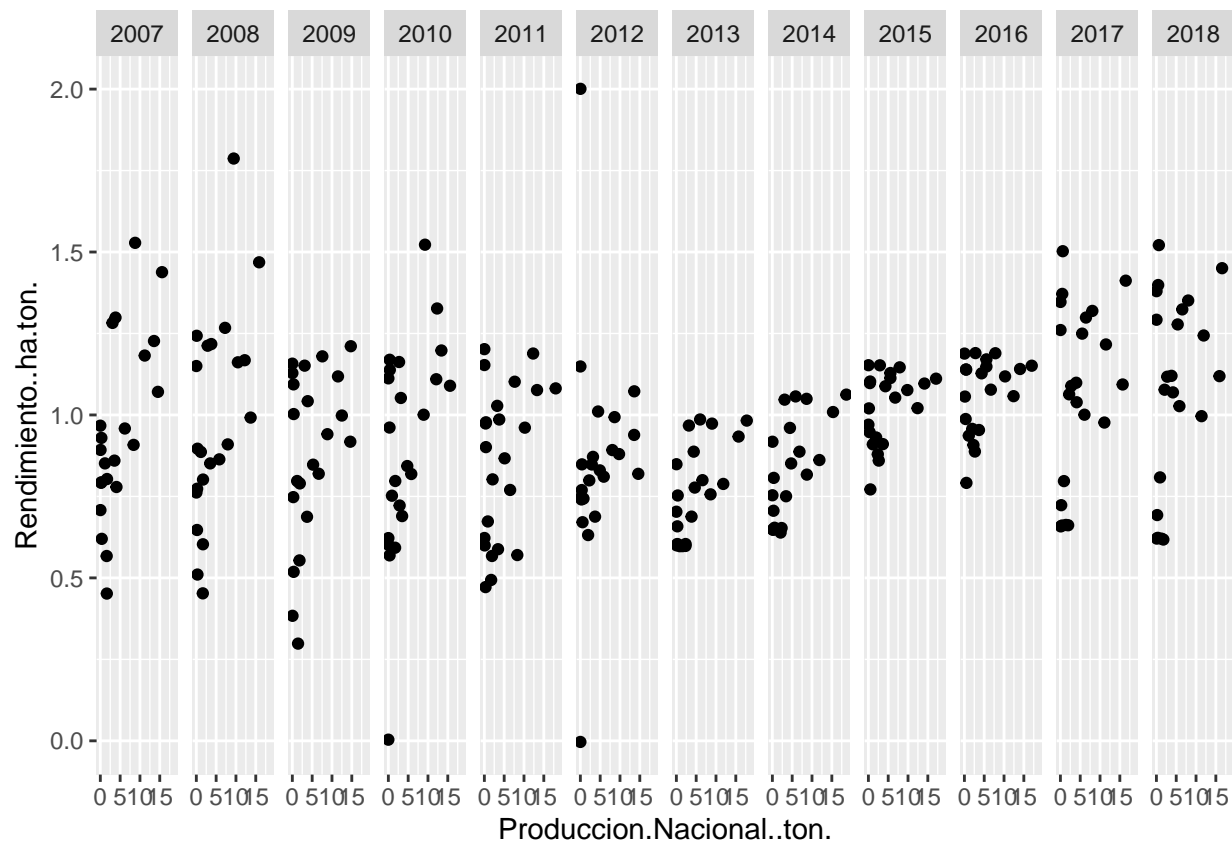
#COMPARANDO CUANDO TIENE REPRESENTADO LOS AÑOS A CONTINUACION

```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio)) +  
  geom_line()
```



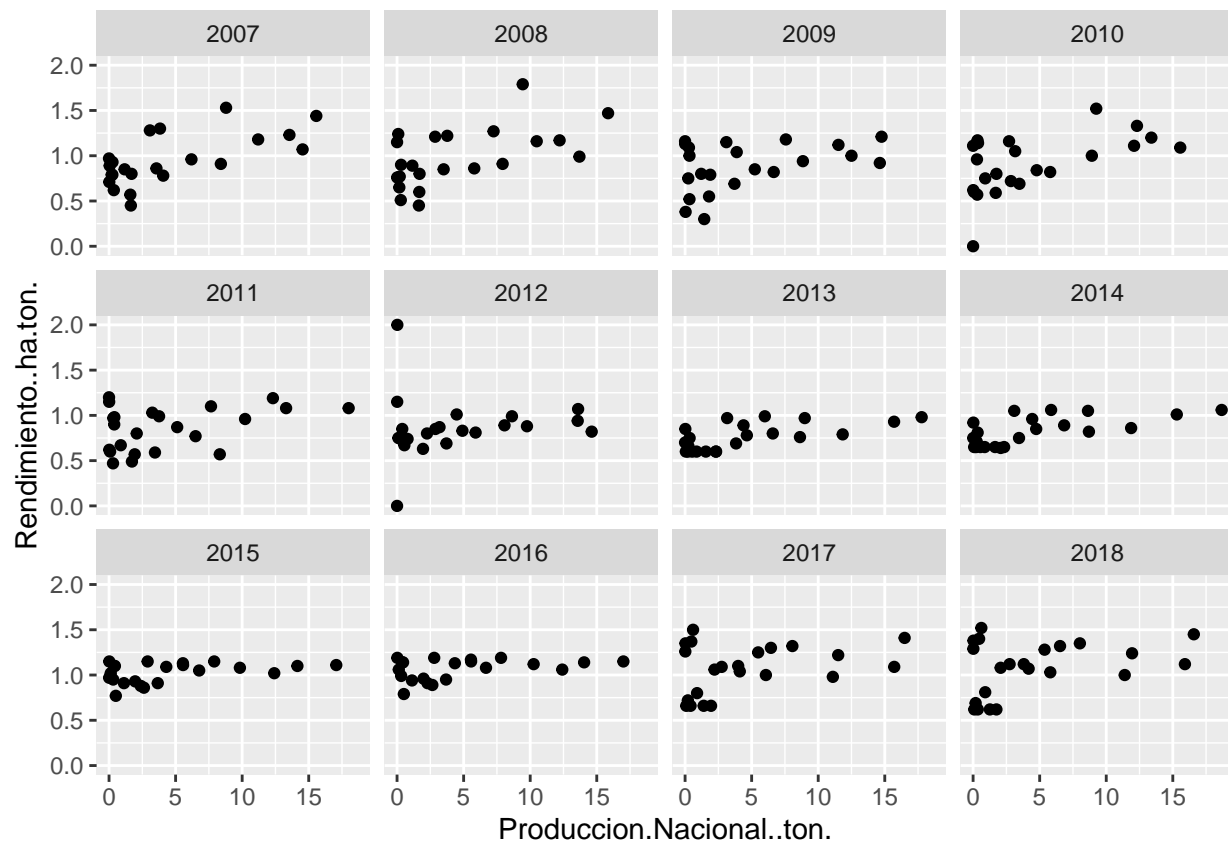
#TAMBIEN SE PUEDE REALIZAR DISPERSION EN PANELES POR AÑOS. #PERMITE ENTENDER O DESCUBRIR PATRONES EN LOS DATOS

```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton.)) +  
  geom_jitter() +  
  facet_grid(.~ Anio)
```

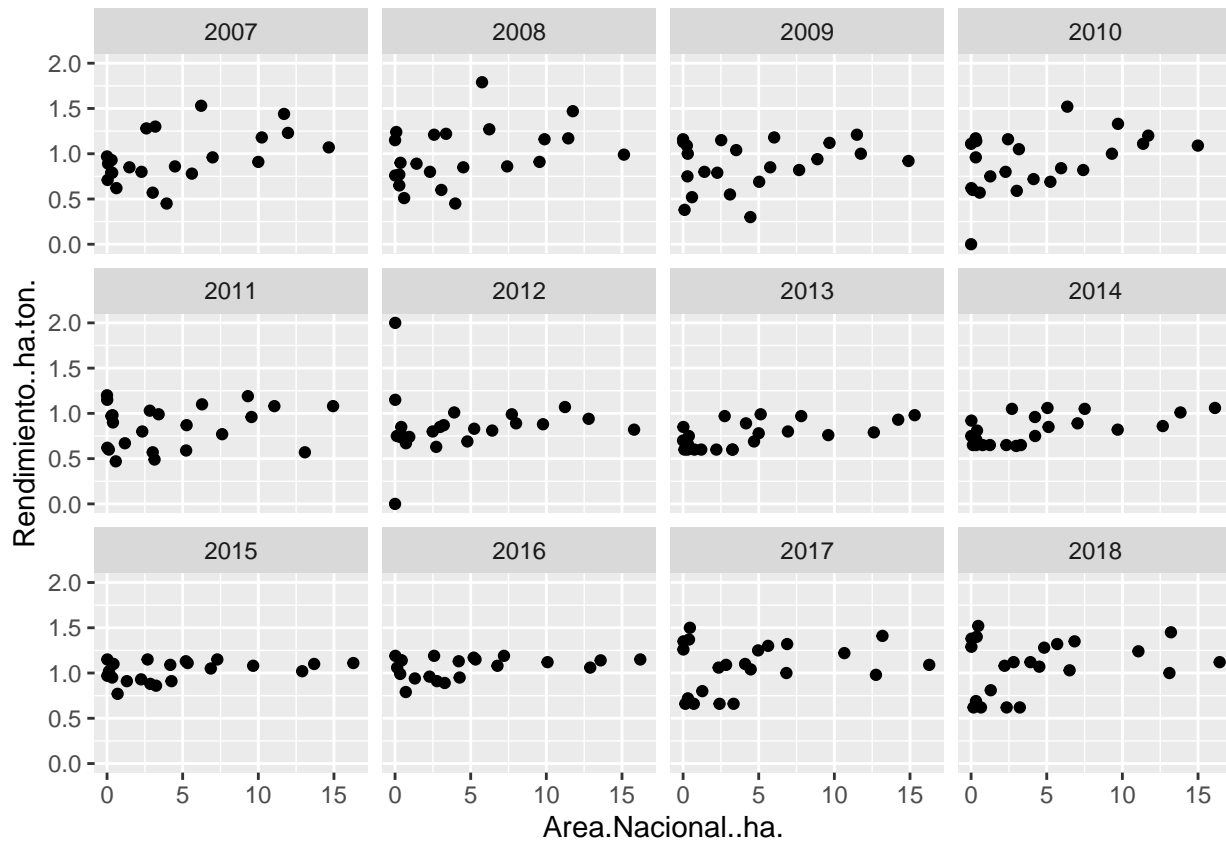


#PANELES SEPARADOS #TAMBIEN SE PUEDE REALIZAR DISPERSION EN PANELES SEPARADOS POR AÑOS

```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton.)) +
  geom_point() +
  facet_wrap(~ Anio)
```



```
ggplot(cape_df, aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton.)) +
  geom_point() +
  facet_wrap(~ Anio)
```

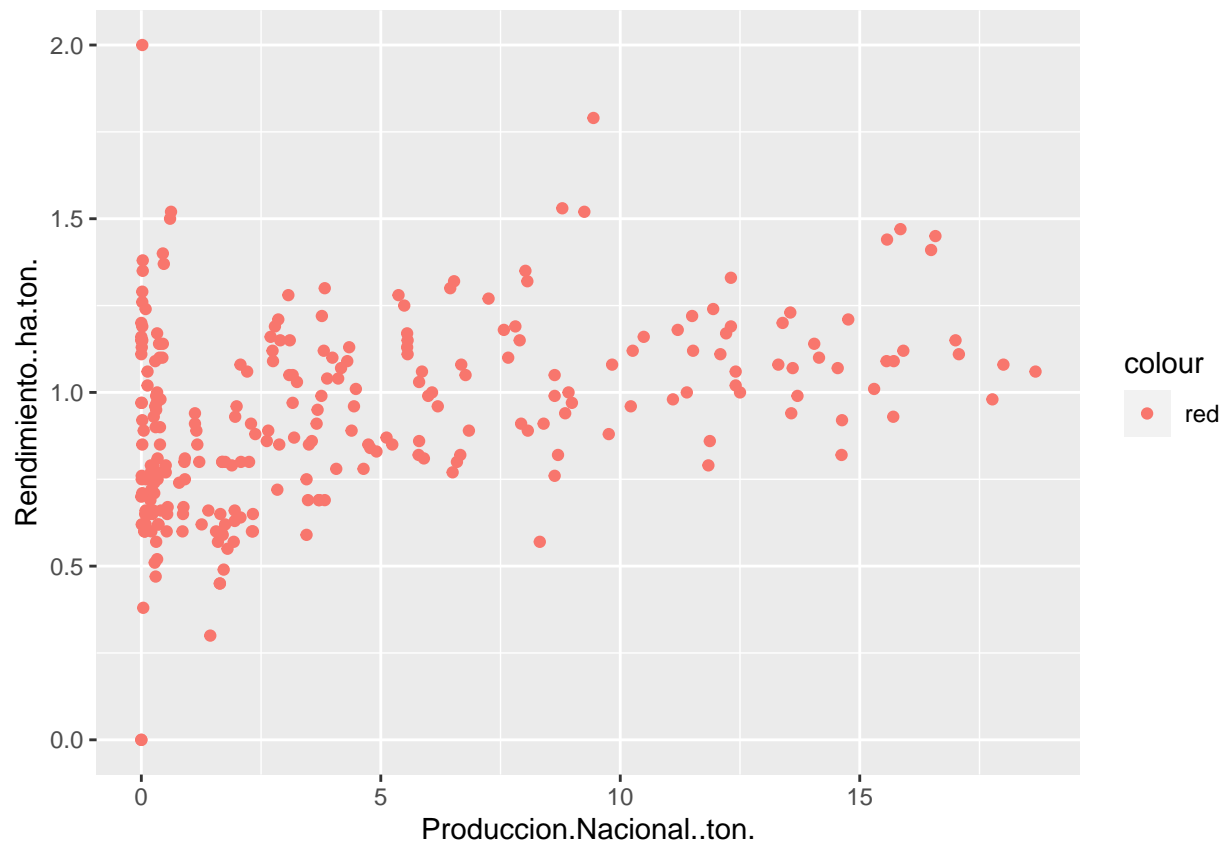


##COMPARANDO EL ANTERIOR CON EL GRAFICO DE DISPERSION SIN AÑOS # OTRA FORMA  
CON GRAFICO DE DISPERSION SIN REPRESENTAR LOS AÑOS

```
data(cafe_df)
```

```
## Warning in data(cafe_df): data set 'cafe_df' not found
```

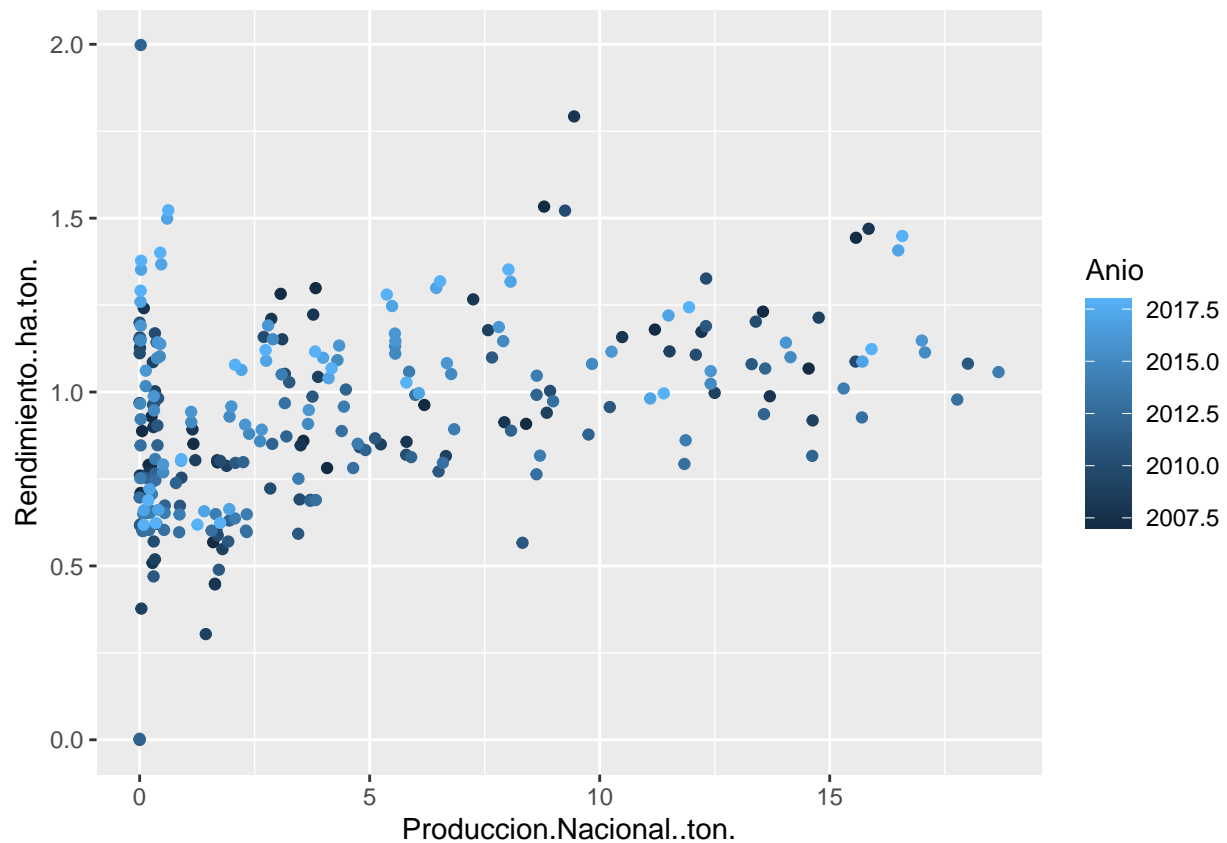
```
ggplot(cafe_df, aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = "red")) +  
  geom_point()
```



#COMPARAR CON GRAFICO DE DISPERSION PRODUCCION NACIONAL VS RENDIMIENTO CON AÑOS

```
ggplot(cafe_df) +  
  geom_jitter(aes(x = Produccion.Nacional..ton., y = Rendimiento..ha.ton., color = Anio))
```

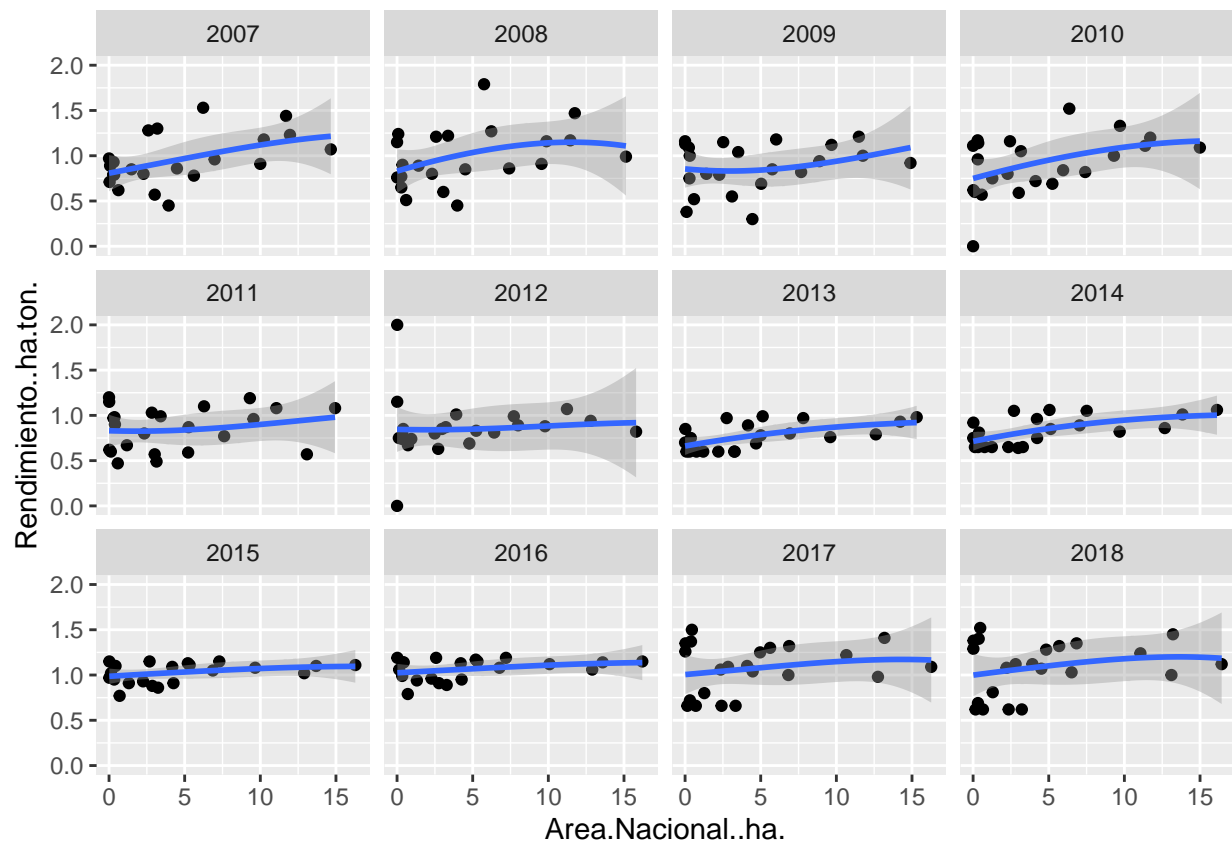




#AHORA SE UTILIZARA UN SUAVIZADOR #LA GRAFICA LINEAL SUAVIZADA

```
ggplot(cafe_df, aes(x = Area.Nacional..ha., y = Rendimiento..ha.ton.)) +
  geom_point() +
  facet_wrap(~ Anio) +
  geom_smooth(span = 3)
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



FIN