

Gráficos con R

Oscar Perpiñán Lamigueiro

Universidad Politécnica de Madrid

① Introducción

② Datos de ejemplo

③ Catálogo de gráficos

Base y grid

- ▶ En R existen dos formas de generar gráficos:
 - ▶ Base graphics
 - ▶ Grid graphics
- ▶ Los gráficos base sólo producen un resultado gráfico, pero no un objeto.
- ▶ Los gráficos grid generan un resultado gráfico **y** un objeto.
- ▶ Dentro del conjunto grid existen dos grandes paquetes: `lattice` y `ggplot2`.

Gráficos lattice

- ▶ Implementación de los gráficos *trellis*, *The Elements of Graphing Data* de Cleveland)
- ▶ Estructura matricial de paneles definida a través de una fórmula.

```
library(lattice)
```

```
xyplot(wt ~ mpg | am, data = mtcars, groups = cyl)
```

- ▶ Documentación: [Código y Figuras del libro](#)

Gráficos ggplot2

- ▶ Implementación de *The Grammar of Graphics* de Wilkinson.
- ▶ Combinación de funciones que proporcionan los componentes (capas) del gráfico.

```
library(ggplot2)

ggplot(mtcars, aes(mpg, wt)) +
  geom_point(aes(colour=factor(cyl))) +
  facet_grid(. ~ am)
```

- ▶ Documentación de ggplot2
- ▶ Código del libro
- ▶ ggplot2 desde lattice (PDF)

① Introducción

② Datos de ejemplo

③ Catálogo de gráficos

Leemos desde el archivo local

```
aranjuez <- read.csv('data/aranjuez.csv')  
  
summary(aranjuez)
```

Añadimos algunas columnas

```
aranjuez$date <- as.Date(aranjuez$X)
```

```
aranjuez$month <- as.numeric(  
  format(aranjuez$date, '%m'))
```

```
aranjuez$year <- as.numeric(  
  format(aranjuez$date, '%Y'))
```

```
aranjuez$day <- as.numeric(  
  format(aranjuez$date, '%j'))
```

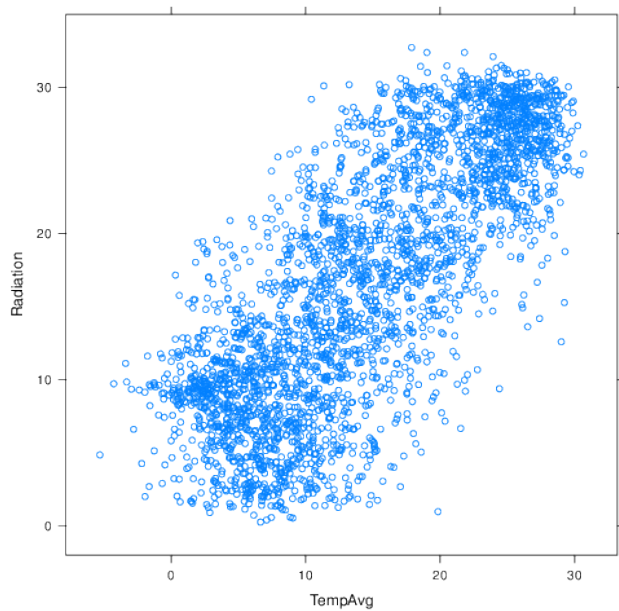
```
aranjuez$quarter <- quarters(aranjuez$date)
```


- ① Introducción
- ② Datos de ejemplo
- ③ Catálogo de gráficos

Gráfico de dispersión de puntos

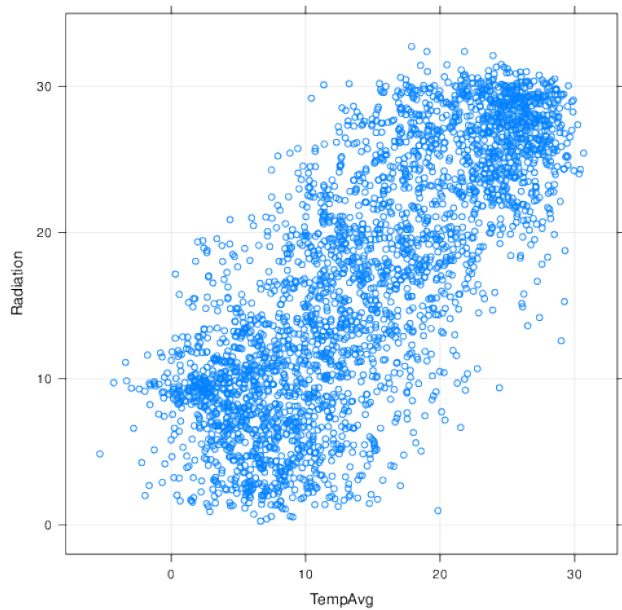
```
xyplot(Radiation ~ TempAvg, data=aranjuez)
```

```
ggplot(aranjuez, aes(TempAvg, Radiation)) +  
  geom_point()
```



Añadimos rejilla

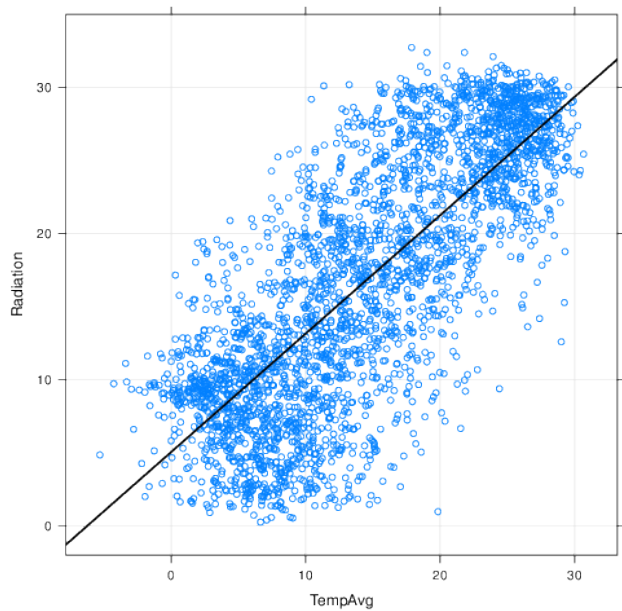
```
xyplot(Radiation ~ TempAvg, data=aranjuez,  
       grid = TRUE)
```



Añadimos regresión lineal

```
xyplot(Radiation ~ TempAvg, data=aranjuez,  
       type=c('p', 'r'), grid = TRUE,  
       lwd=2, col.line='black')
```

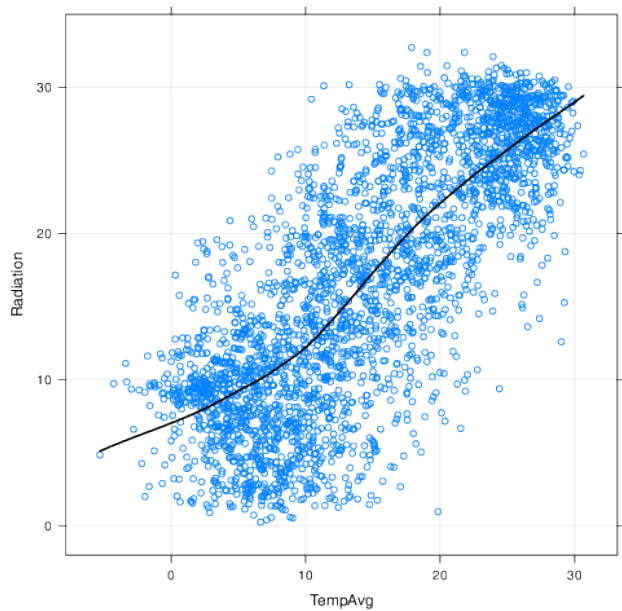
```
ggplot(aranjuez, aes(TempAvg, Radiation)) +  
  geom_point() +  
  geom_smooth(method = "lm")
```



Añadimos ajuste local

```
xyplot(Radiation ~ TempAvg, data=aranjuez,  
       type=c('p', 'smooth'), grid = TRUE,  
       lwd=2, col.line='black')
```

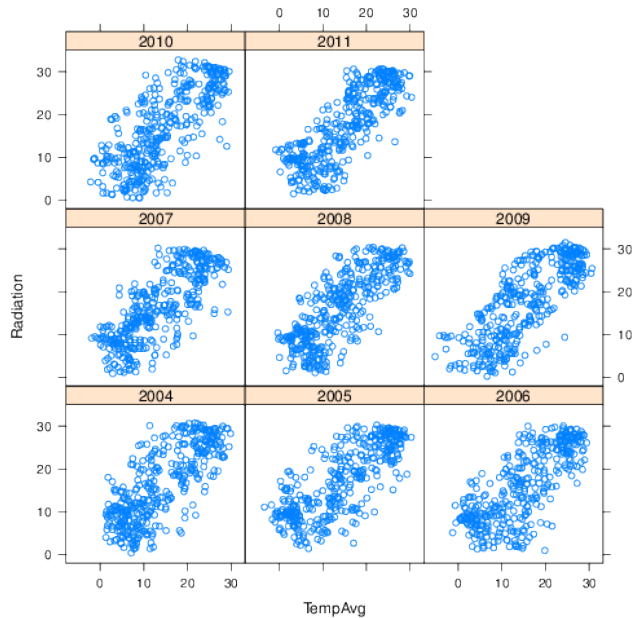
```
ggplot(aranjuez, aes(TempAvg, Radiation)) +  
  geom_point() +  
  geom_smooth()
```

Paneles

```
xyplot(Radiation ~ TempAvg|factor(year),  
       data=aranjuez)
```

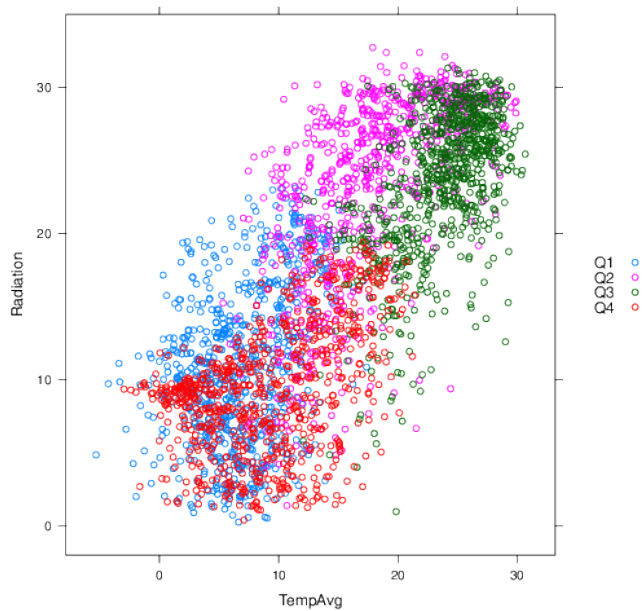
```
ggplot(aranjuez, aes(TempAvg, Radiation)) +  
  geom_point() +  
  facet_wrap(~factor(year))
```



Grupos

```
xyplot(Radiation ~ TempAvg, groups=quarter,  
       data=aranjuez, auto.key=list(space='right'))
```

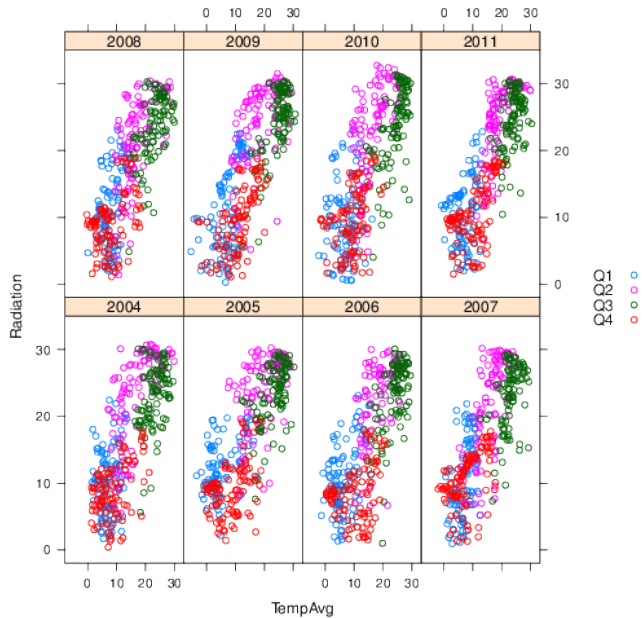
```
ggplot(aranjuez, aes(TempAvg, Radiation,  
                     color = quarter)) +  
  geom_point()
```



Paneles y grupos

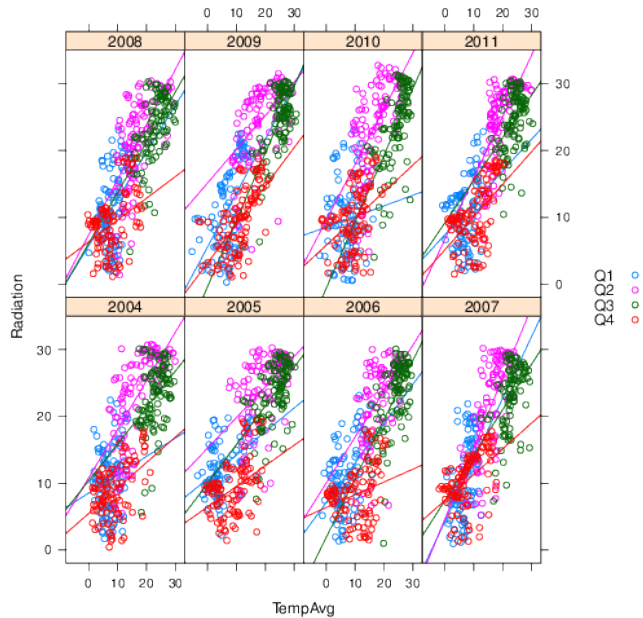
```
xyplot(Radiation ~ TempAvg|factor(year),  
       groups=quarter,  
       data=aranjuez,  
       layout=c(4, 2),  
       auto.key=list(space='right'))
```

```
ggplot(aranjuez, aes(TempAvg, Radiation,  
                     color = quarter)) +  
  geom_point() +  
  facet_wrap(~factor(year))
```



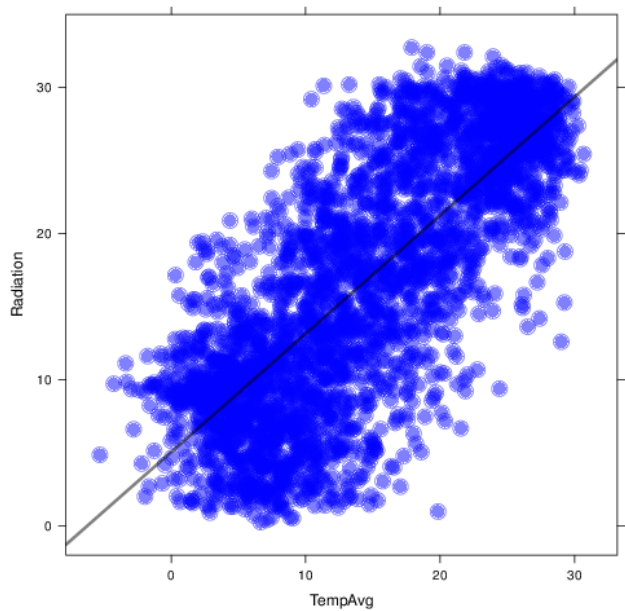
Paneles y grupos

```
xyplot(Radiation ~ TempAvg|factor(year),  
       groups=quarter,  
       data=aranjuez,  
       layout=c(4, 2),  
       type=c('p', 'r'),  
       auto.key=list(space='right'))
```

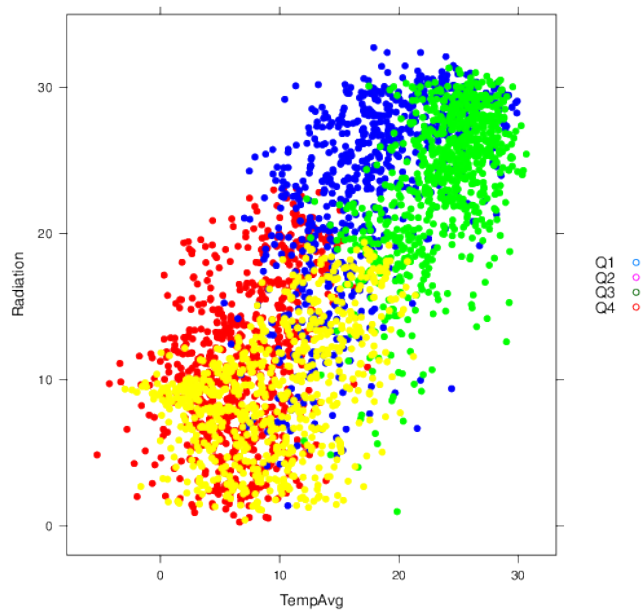
Colores y tamaños

```
xyplot(Radiation ~ TempAvg,  
       type=c('p', 'r'),  
       cex=2, col='blue',  
       alpha=.5, pch=19,  
       lwd=3, col.line='black',  
       data=aranjuez)
```



Colores con grupos

```
xyplot(Radiation ~ TempAvg,  
       group=quarter,  
       col=c('red', 'blue', 'green', 'yellow'),  
       pch=19,  
       auto.key=list(space='right'),  
       data=aranjuez)
```



Colores con grupos: `par.settings` y `simpleTheme`

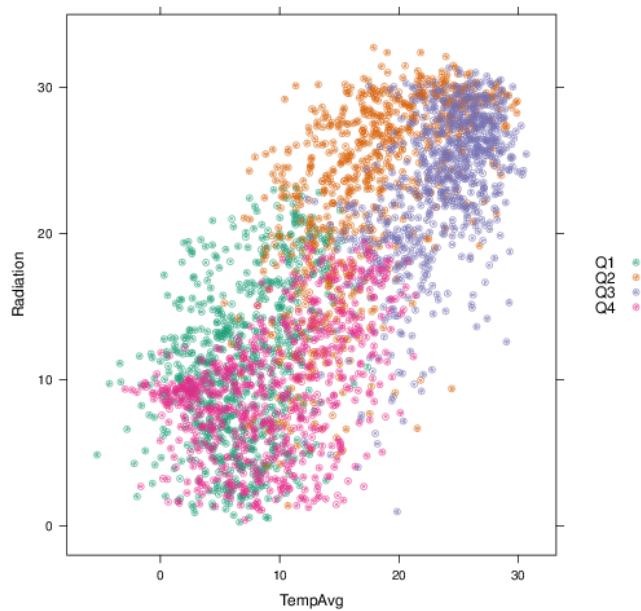
- Primero definimos el tema con `simpleTheme`

```
myTheme <- simpleTheme(col=c('red', 'blue',  
                             'green', 'yellow'),  
                       pch=19, alpha=.6)
```

Colores con grupos: `par.settings` y `simpleTheme`

- Aplicamos el resultado en `par.settings`

```
xyplot(Radiation ~ TempAvg,  
       groups=quarter,  
       par.settings=myTheme,  
       auto.key=list(space='right'),  
       data=aranjuez)
```



Colores: brewer.pal

```
library(RColorBrewer)

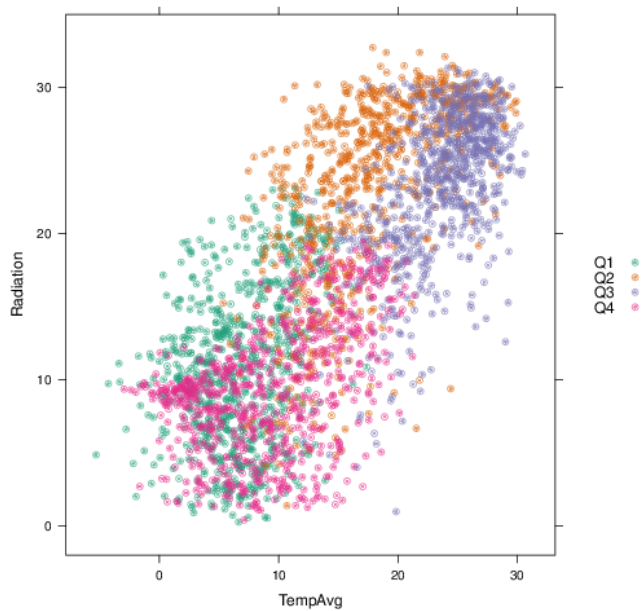
myPal <- brewer.pal(n = 4, 'Dark2')

myTheme <- simpleTheme(col = myPal,
                       pch=19, alpha=.6)
```

ColorBrewer: <http://colorbrewer2.org/>

Asignamos paleta con `par.settings`

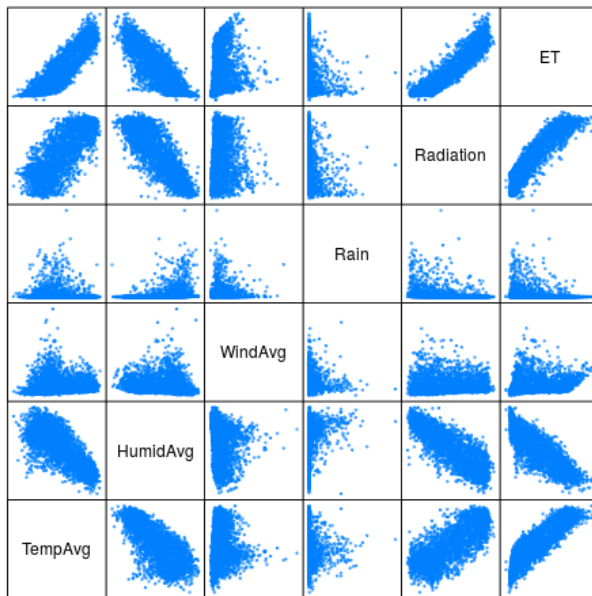
```
xyplot(Radiation ~ TempAvg,  
       groups=quarter,  
       par.settings=myTheme,  
       auto.key=list(space='right'),  
       data=aranjuez)
```



Matriz de gráficos de dispersión

```
splom(aranjuez[,c("TempAvg", "HumidAvg", "WindAvg",  
                  "Rain", "Radiation", "ET")],  
      pscale=0, alpha=0.6, cex=0.3, pch=19)
```

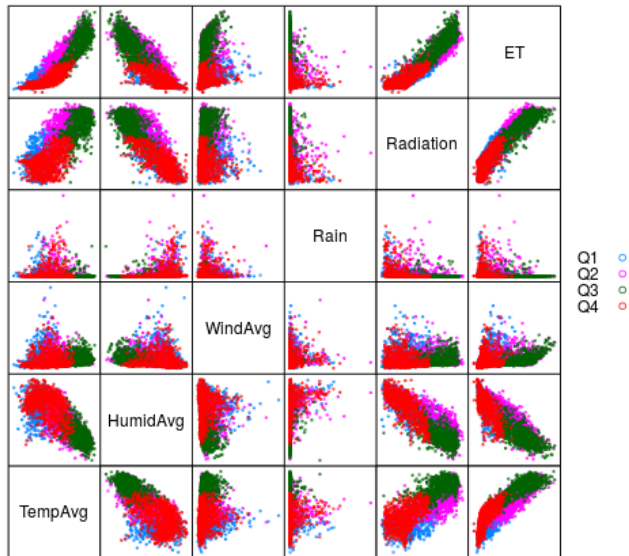
```
library(GGally)  
ggpairs(aranjuez)
```



Scatter Plot Matrix

Matriz de gráficos de dispersión

```
splom(aranjuez[,c("TempAvg", "HumidAvg", "WindAvg",  
                  "Rain", "Radiation", "ET")],  
      groups=aranjuez$quarter,  
      auto.key=list(space='right'),  
      pscale=0, alpha=0.6, cex=0.3, pch=19)
```

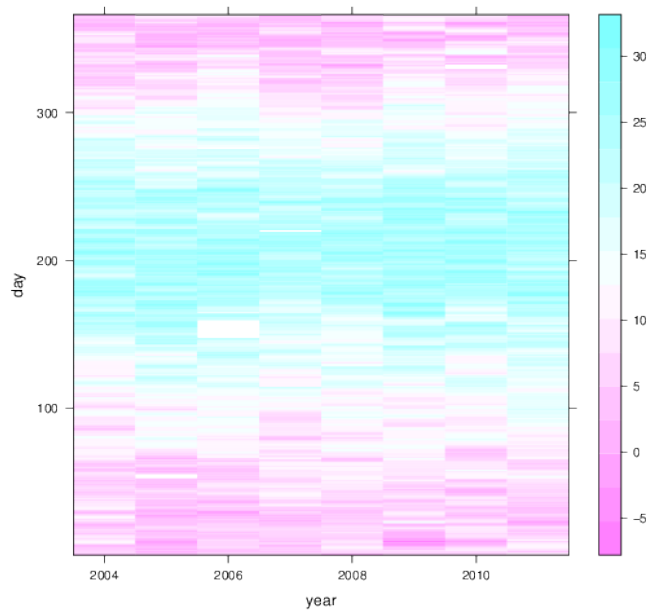


Scatter Plot Matrix

Mapa de niveles

```
levelplot(TempAvg ~ year * day, data = aranjuez)
```

```
ggplot(aranjuez, aes(year, day)) +  
  geom_raster(aes(fill = TempAvg))
```

levelplot con una paleta mejor

- ▶ Usamos `colorRampPalette` para generar una función que interpola colores a partir de una paleta

```
levelPal <- colorRampPalette(  
  brewer.pal(n = 9, 'Oranges'))
```

- ▶ Comprobamos que es una función generadora de colores

```
levelPal(14)
```

- ▶ Usamos esta función con `col.regions`

```
levelplot(TempAvg ~ year * day,  
  col.regions = levelPal,  
  data = aranjuez)
```

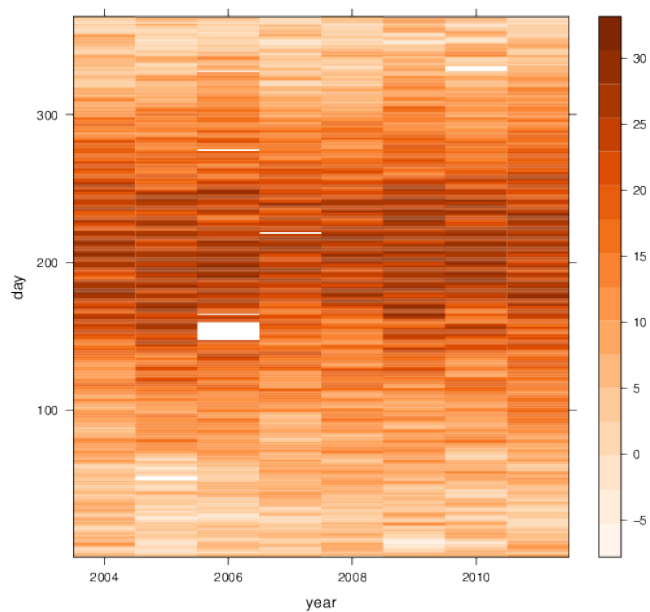
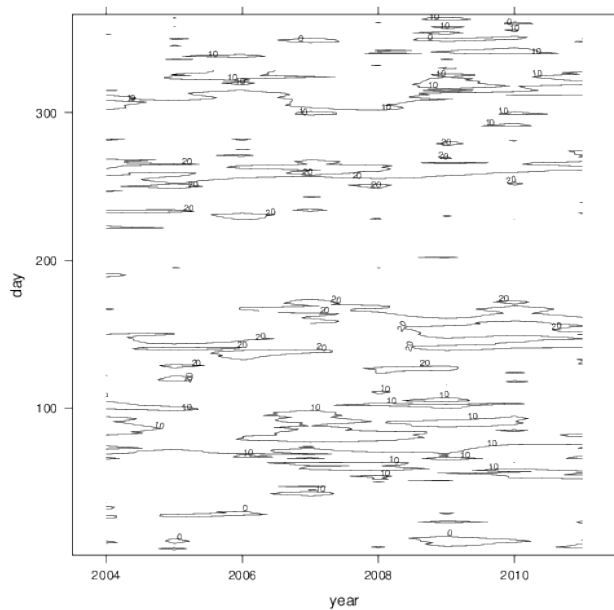


Gráfico de contornos

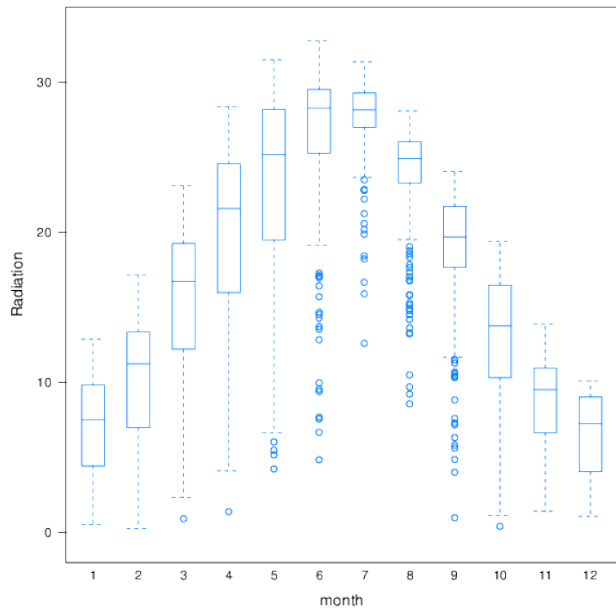
```
contourplot(TempAvg ~ year * day,  
  data = aranjuez,  
  lwd = .5,  
  labels = list(cex = 0.6),  
  label.style = 'align',  
  cuts = 5)
```



Box-and-Whiskers

```
bwplot(Radiation ~ month, data=aranjuez,  
       horizontal = FALSE, pch='|')
```

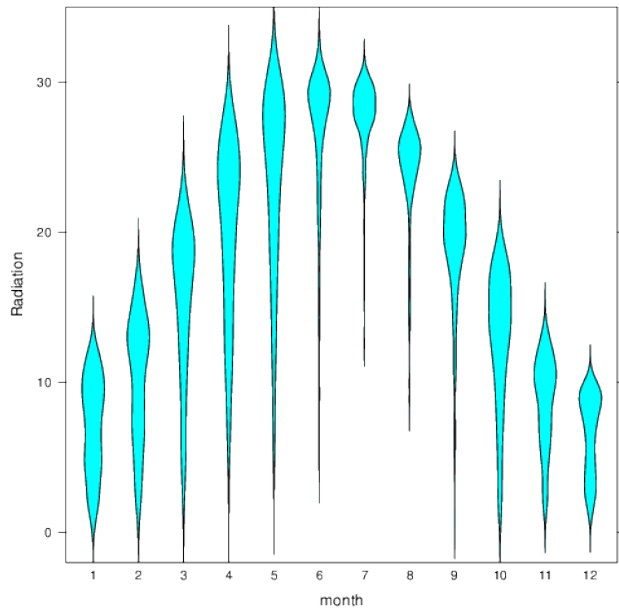
```
ggplot(aranjuez, aes(factor(month), Radiation)) +  
  geom_boxplot()
```



Box-and-Whiskers

```
bwplot(Radiation ~ month, data=aranjuez,  
       horizontal=FALSE,  
       panel=panel.violin)
```

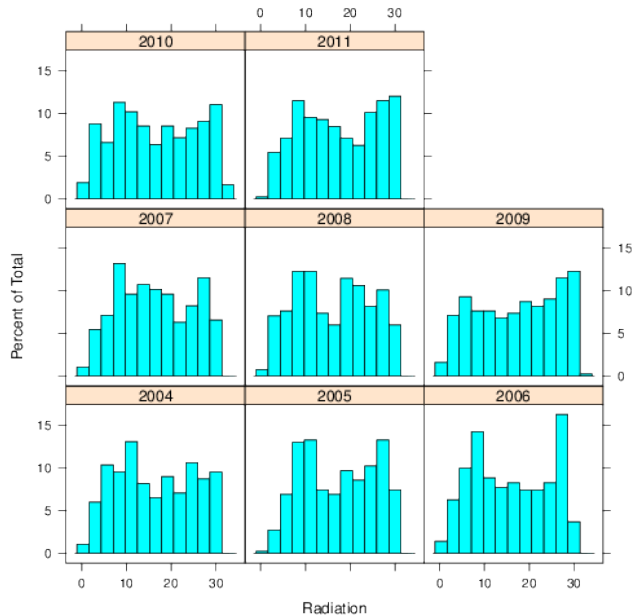
```
ggplot(aranjuez, aes(factor(month), Radiation)) +  
  geom_violin()
```

Histogramas

```
histogram(~ Radiation|factor(year), data=aranjuez)
```

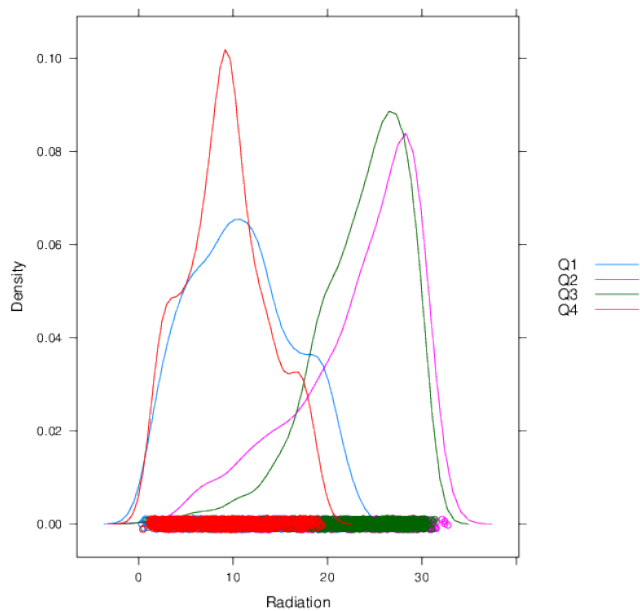
```
ggplot(aranjuez, aes(Radiation)) +  
  geom_histogram() +  
  facet_wrap(~factor(year))
```



Gráficos de densidad

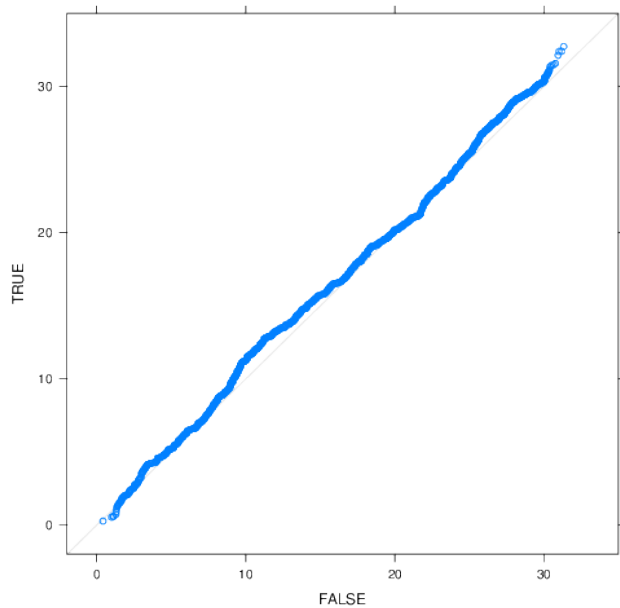
```
densityplot(~ Radiation, groups=quarter,  
            data=aranjuez,  
            auto.key=list(space='right'))
```

```
ggplot(aranjuez, aes(Radiation, color = quarter)) +  
  geom_density()
```



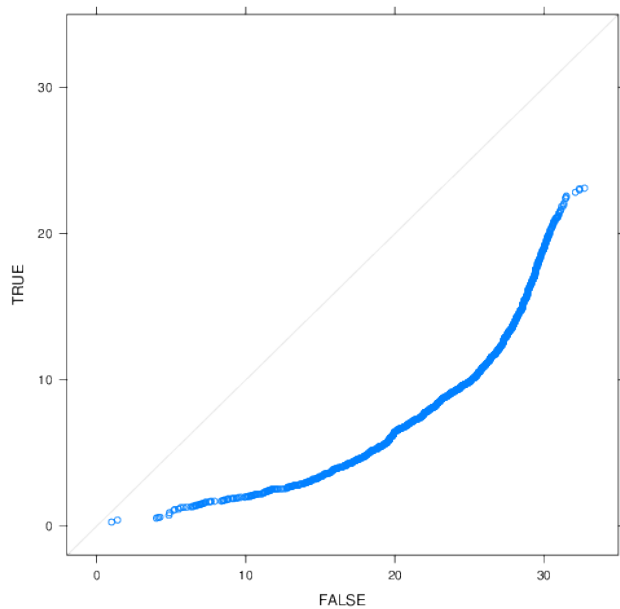
Quantile-Quantile

```
firstHalf <- aranjuez$quarter %in% c('Q1', 'Q2')  
  
qq(firstHalf ~ Radiation, data=aranjuez)
```



Quantile-quantile

```
winter <- aranjuez$quarter %in% c('Q1', 'Q4')  
qq(winter ~ Radiation, data=aranjuez)
```

Quantile-Quantile

```
qqmath(~TempAvg, data=aranjuez,  
       groups=year, distribution=qnorm)
```

