CargarDatos

Ignacio Acosta, Mauro Loprete y Sofía Itté

04 de June

Carga de datos

```
Datos <- read.table(
    "UScrime.txt",
    header = TRUE,
    dec = ","
)
```

Primera idea: Incluir todas las variables :

```
lm(
    y ~ . ,
    data = Datos
) %>% summary()
```

```
##
## Call:
## lm(formula = y ~ ., data = Datos)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -395.74  -98.09   -6.69  112.99  512.67
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5984.2876  1628.3184  -3.675 0.000893 ***
## M               8.7830     4.1714   2.106 0.043443 *
## So             -3.8035   148.7551  -0.026 0.979765
## Ed             18.8324     6.2088   3.033 0.004861 **
## Po1            19.2804    10.6110   1.817 0.078892 .
## Po2           -10.9422    11.7478  -0.931 0.358830
## LF             -0.6638     1.4697  -0.452 0.654654
## M.F             1.7407     2.0354   0.855 0.398995
## Pop            -0.7330     1.2896  -0.568 0.573845
## NW              0.4204     0.6481   0.649 0.521279
## U1             -5.8271     4.2103  -1.384 0.176238
```

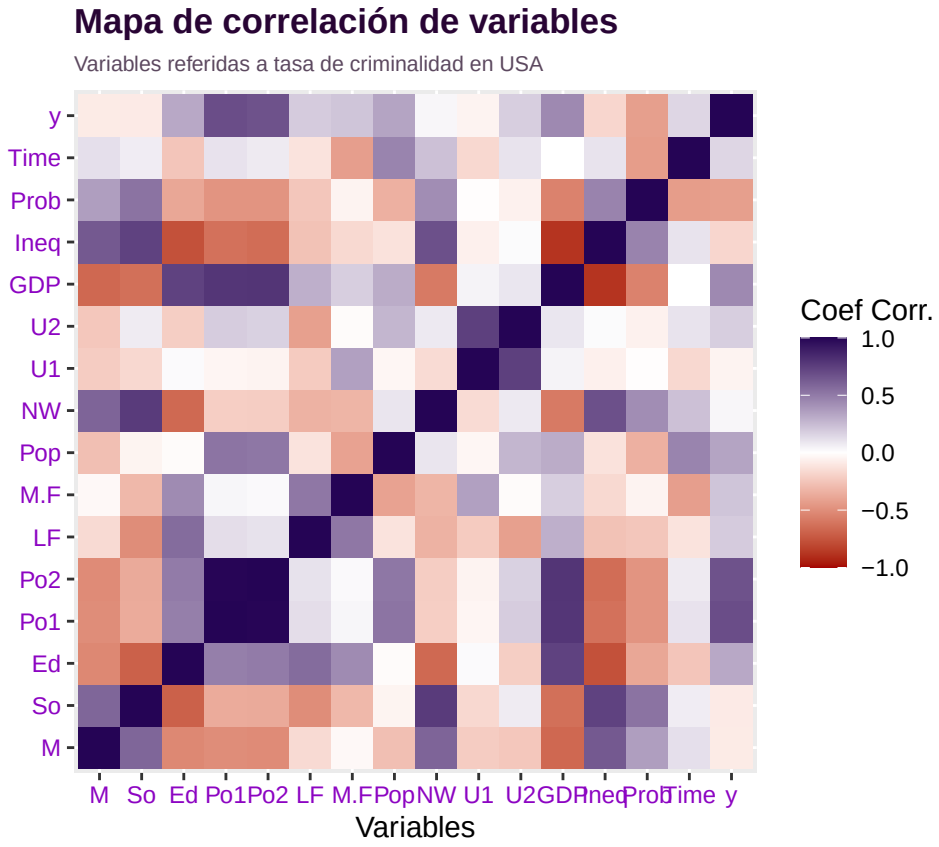
```
## U2             16.7800     8.2336   2.038 0.050161 .
## GDP             0.9617     1.0367   0.928 0.360754
## Ineq            7.0672     2.2717   3.111 0.003983 **
## Prob        -4855.2658  2272.3746  -2.137 0.040627 *
## Time           -3.4790     7.1653  -0.486 0.630708
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 209.1 on 31 degrees of freedom
## Multiple R-squared:  0.8031, Adjusted R-squared:  0.7078
## F-statistic: 8.429 on 15 and 31 DF,  p-value: 3.539e-07
```

En este caso se puede ver que pocas variables son significativas, pero el modelo en general si lo es ...

Pasos siguientes :

- Hay que empezar el proceso de ... (Backward/Forward/Stepwise)
- Diagnostico ...

```
qplot(x=Var1,
      y=Var2,
      data = melt(cor(Datos, use = "p")),
      fill = value,
      geom="tile"
)+scale_fill_gradient2(limits = c(-1, 1),
                       low = "#A50303", high = "#250455")+
  theme(aspect.ratio = 1,
        plot.title = element_text(
          face = "bold",
          color = "#280434FF"),
        plot.subtitle = element_text(
          size=8,
          color="#5C485F"
        ),
        axis.text.x = element_text(
          color = "#8C04C2"
        ),
        axis.text.y = element_text(
          color="#8C04C2"
        )
        )+
  labs(title="Mapa de correlación de variables",
       subtitle = "Variables referidas a tasa de criminalidad en USA",
       x="Variables",
       y="",
       fill="Coef Corr.")
```


Mapa de correlación de variables
Variables referidas a tasa de criminalidad en USA
y
Time
Prob
Ineq
GDP
U2
U1
NW
Pop
M.F
LF
Po2
Po1
Ed
So
M
M So Ed Po1 Po2 LF M.F Pop NW U1 U2 GDP Ineq Prob Time y
Variables
Coef Corr.
1.0
0.5
0.0
−0.5
−1.0