

Ejercicio-04-12-2019-Regression

Daniel Ferreira Zanchetta y Lais Silva Almeida Zanchetta

Fecha de Entrega: 11/12/2019

Exercises: Regression

1. Enumere cuales son las hipótesis que asumimos al hacer una regresión múltiple entre una variable de respuesta y unas variables predictoras, x_1 , ..., x_p .

Resp.: Podemos asumir hipótesis sin hacer suposiciones, donde se hace una regresión local estimando valores de r (fit) y error (e) desde los datos. O, por otro lado, haciendo suposiciones. En esta alternativa, tenemos que buscar la función de r (fit) asumiendo un valor de B_0 (intercept) y $B_n X_n$ (slope de las n predictoras - o mejor, la pendiente). El B_0 sirve para ajustar la nube de puntos, mientras que los predictores $B_n X_n$ son valores teóricos que indican la variación de B_0 (y) por cada incremento de X .

2. En un modelo de regresión, como se calcula y como se interpreta el coeficiente de determinación R^2 .

Resp.: Es calculado a través de la suma de la variación cuadrática de y explicada por el modelo (a través de los predictores) y el valor residual cuadrático. Cuanto más cerca de 1 se sitúe su valor, mayor será el ajuste del modelo a la variable que estamos intentando explicar. Esto se calcula a través del coseno cuadrado del valor de la suma total de cuadrados (TSS) y del valor de la variación explicada por el modelo (ESS).

3. Lea el fichero "BCN_pisos.txt". Del fichero resultante seleccione 2/3 partes como muestra de training y la tercera parte restante como muestra test.

```
bcnpisos<-  
read.csv("C:/Users/dev/Downloads/Sessio3_cursBigData_LinearModel/exer3/bc  
n_pisos.txt", header=TRUE, sep="\t")
```

Inicialmente hemos mirado el dataframe BCN Pisos y habiamos planteado dividir 2/3 a través de la información del Ascensor, donde el 2/3 sería para CON ascensor, y el 1/3 SIN ascensor. Sin embargo, por fin, hemos decidido utilizar una función para definir los datos de training y de lo test.

```
library(caret)
```

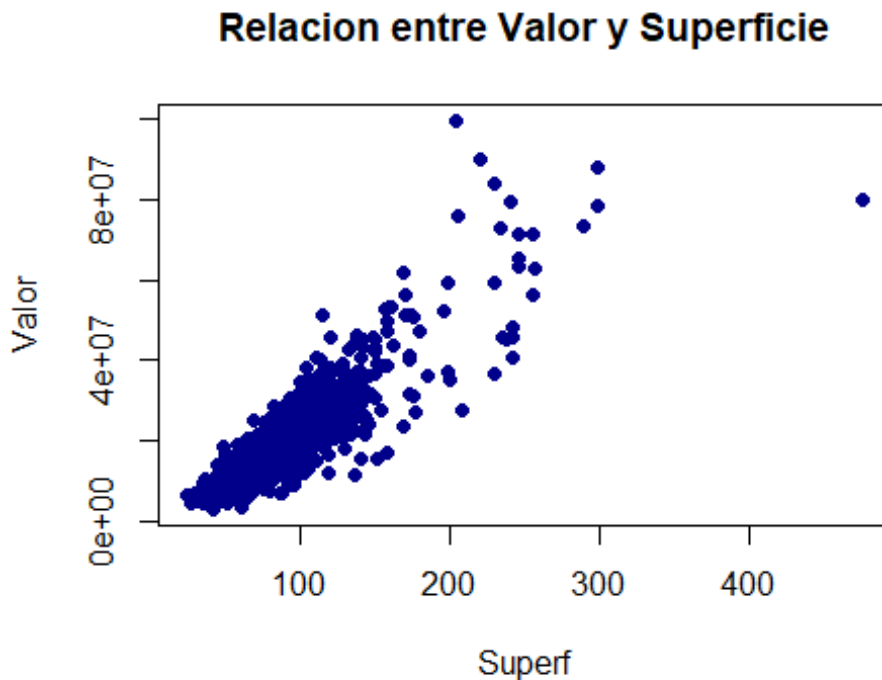
```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

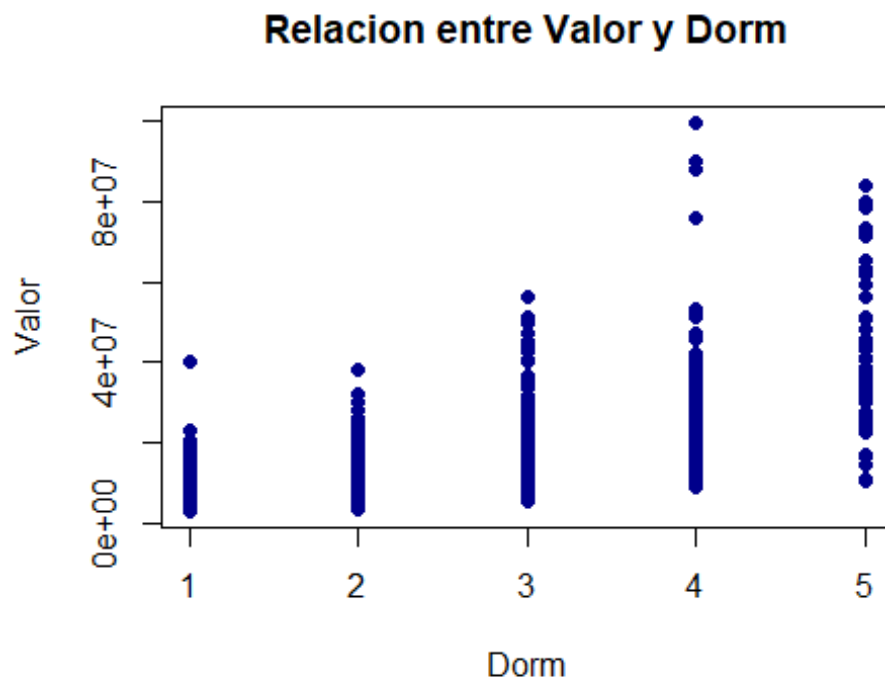
```
inTraining <- createDataPartition(bcnpisos$Superf,times=1,p=0.75,
list=FALSE)
training_bcnpisos <- bcnpisos[inTraining,]
test_bcnpisos <- bcnpisos[-inTraining,]
```

4. Con la muestra de training, efectúe la representación gráfica de la variable “Valor” respecto del resto de variables del fichero. Calcule la correlación entre la variable “Valor” y el resto de variables numéricas.

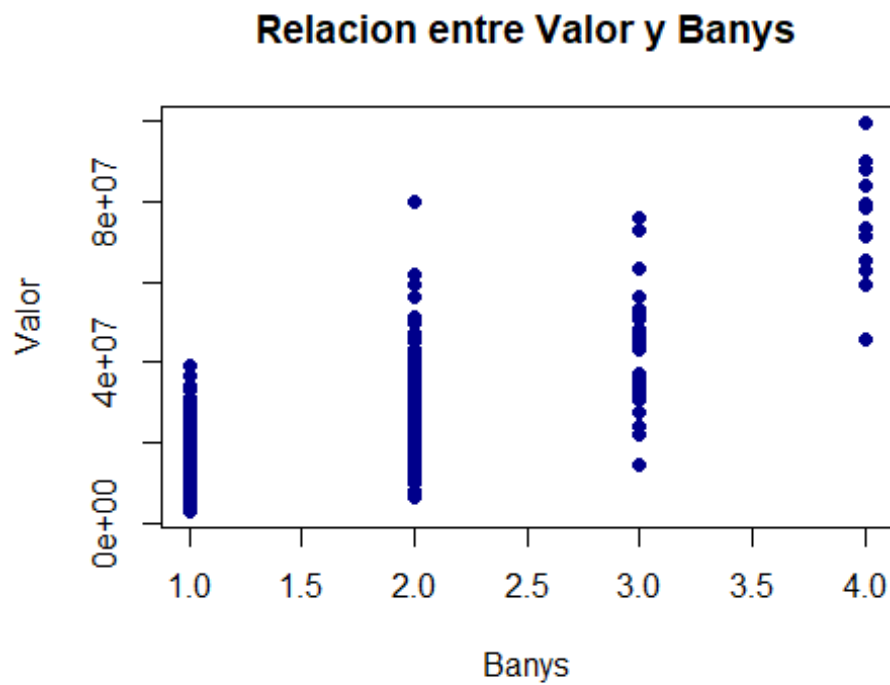
```
# Valor con relación a Superf
plot(training_bcnpisos$Superf, training_bcnpisos$Valor,xlab = "Superf",
ylab = "Valor", main="Relacion entre Valor y Superficie", type = "p",
pch=19,col="darkblue")
```



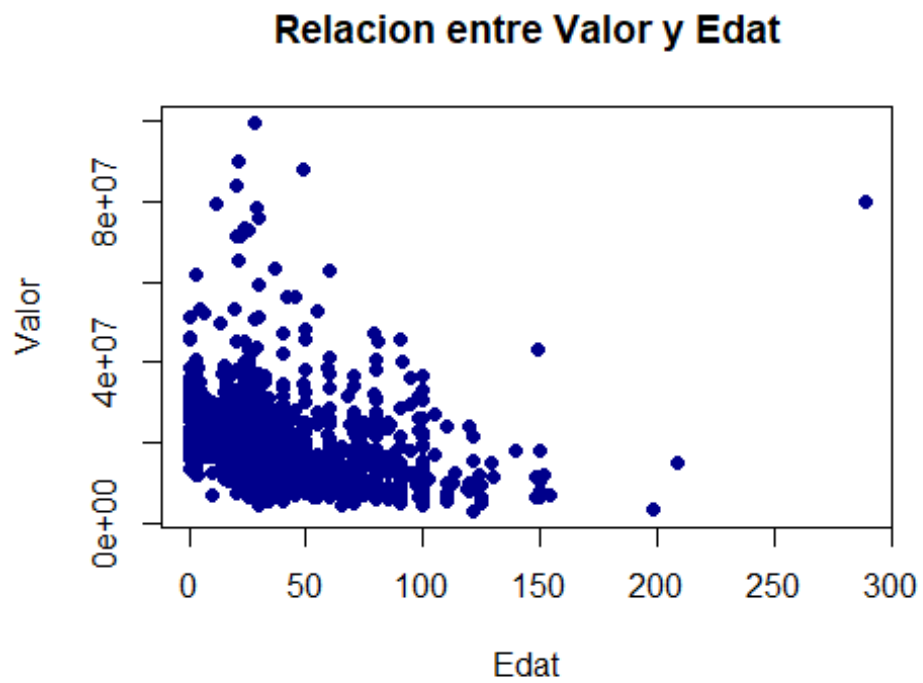
```
# Valor con relación a Dorm
plot(training_bcnpisos$Dorm, training_bcnpisos$Valor,xlab = "Dorm", ylab = "Valor", main="Relacion entre Valor y Dorm", type = "p",
pch=19,col="darkblue")
```



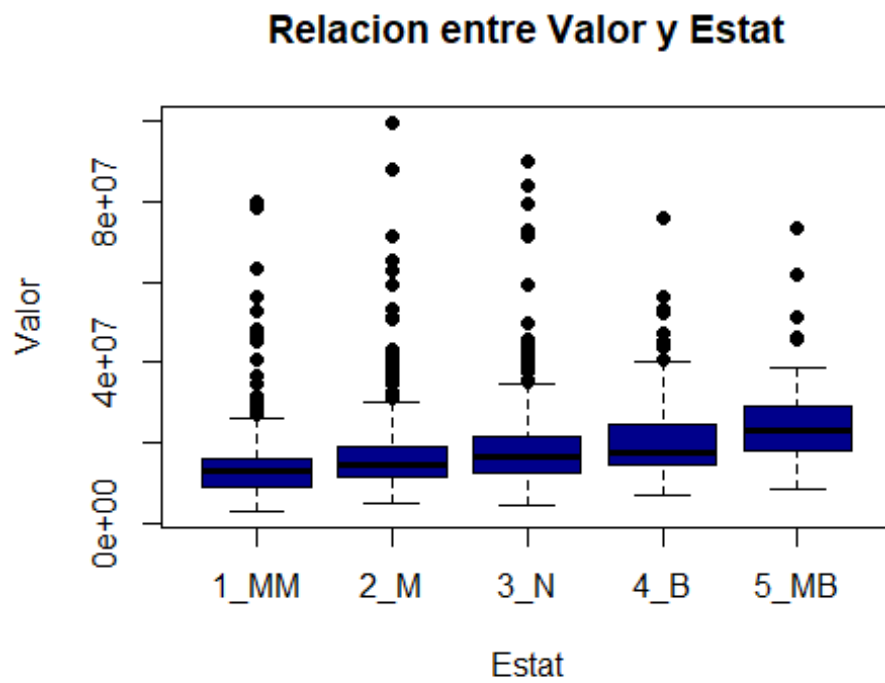
```
# Valor con relación a Banys  
plot(training_bcnpisos$Banys, training_bcnpisos$Valor, xlab = "Banys",  
      ylab = "Valor", main = "Relacion entre Valor y Banys", type = "p",  
      pch = 19, col = "darkblue")
```



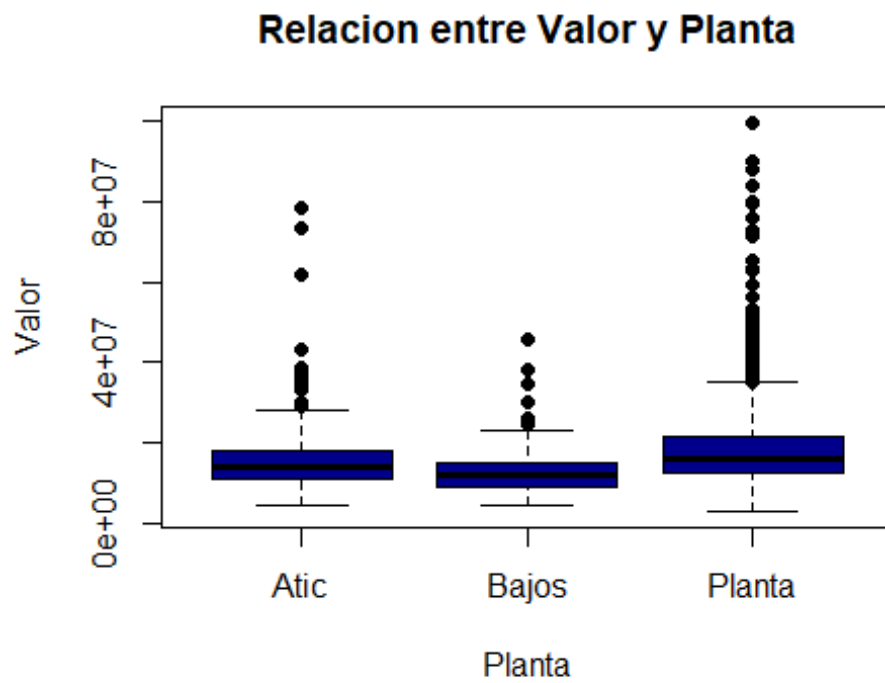
```
# Valor con relación a Edat
plot(training_bcnpisos$Edat, training_bcnpisos$Valor, xlab = "Edat", ylab = "Valor", main="Relacion entre Valor y Edat", type = "p", pch=19, col="darkblue")
```



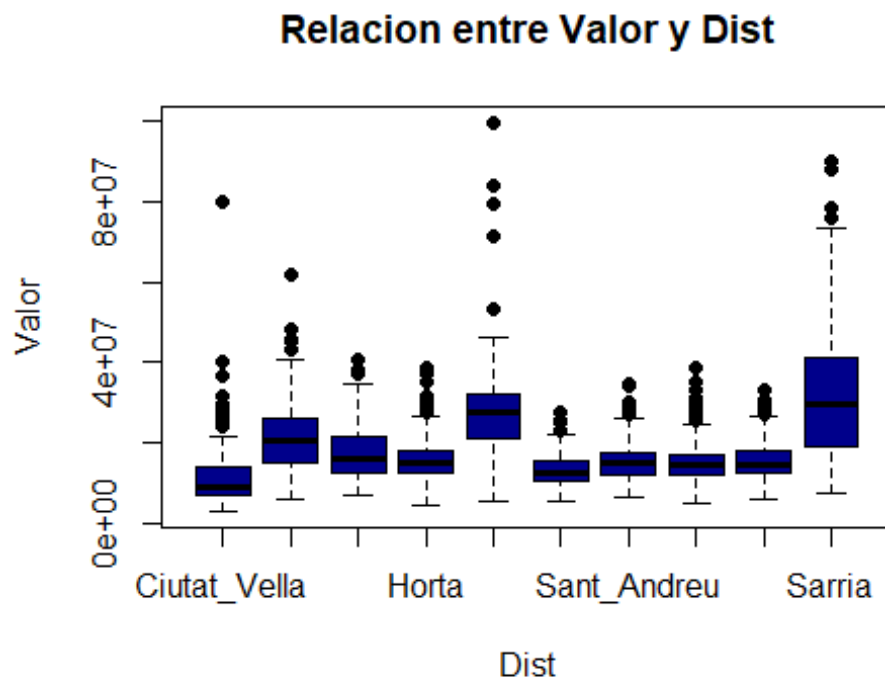
```
# Valor con relación a Estat  
plot(training_bcnpisos$Estat, training_bcnpisos$Valor, xlab = "Estat",  
      ylab = "Valor", main = "Relacion entre Valor y Estat", type = "p",  
      pch = 19, col = "darkblue")
```



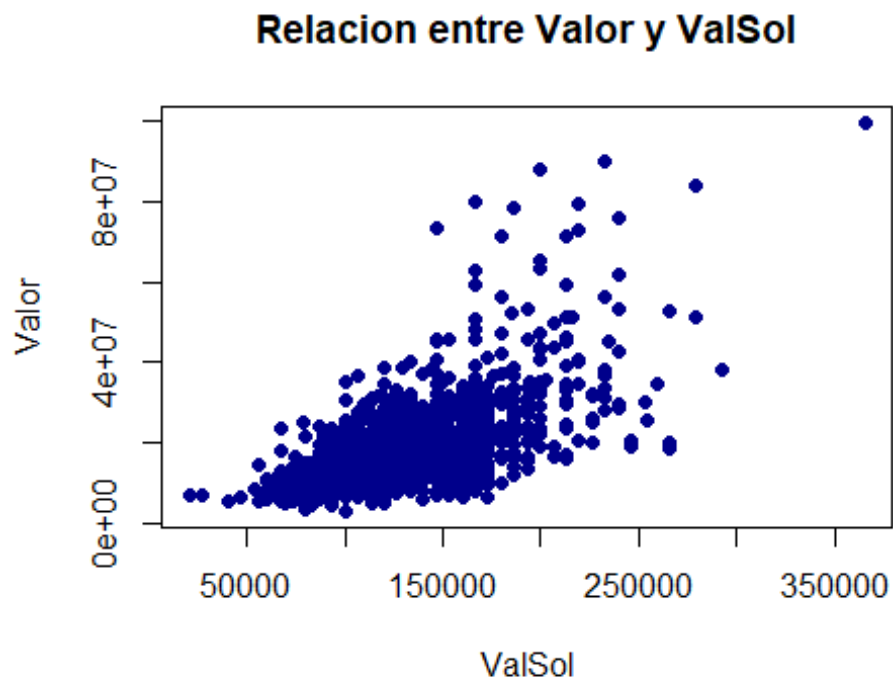
```
# Valor con relación a Planta
plot(training_bcnpisos$Planta, training_bcnpisos$Valor, xlab = "Planta",
      ylab = "Valor", main="Relacion entre Valor y Planta", type = "p",
      pch=19,col="darkblue")
```



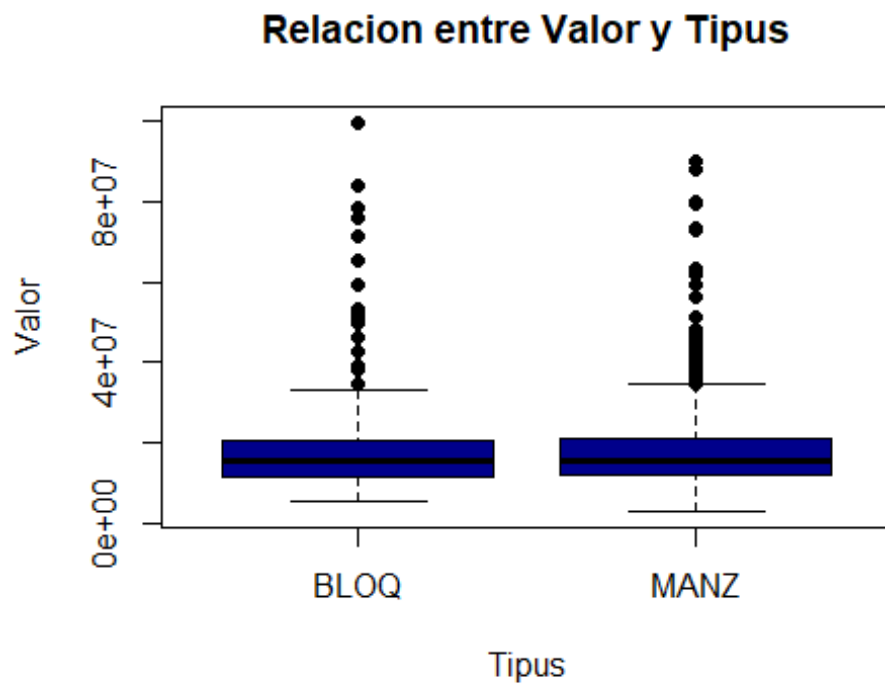
```
# Valor con relación a Dist  
plot(training_bcnpisos$Dist, training_bcnpisos$Valor,xlab = "Dist", ylab  
= "Valor", main="Relacion entre Valor y Dist", type = "p",  
pch=19,col="darkblue")
```



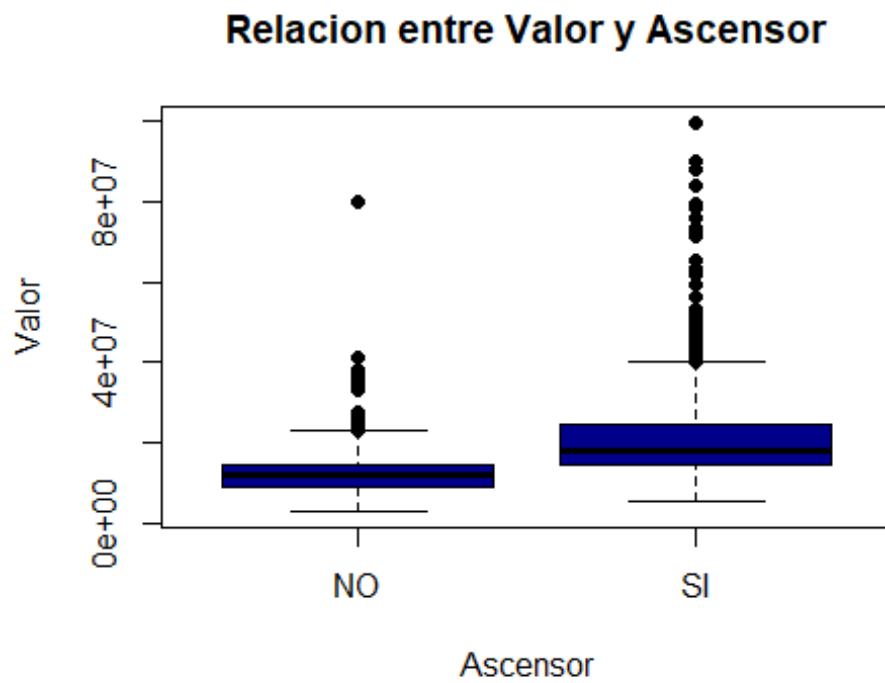
```
# Valor con relación a ValSol
plot(training_bcnpisos$ValSol, training_bcnpisos$Valor, xlab = "ValSol",
      ylab = "Valor", main="Relacion entre Valor y ValSol", type = "p",
      pch=19, col="darkblue")
```

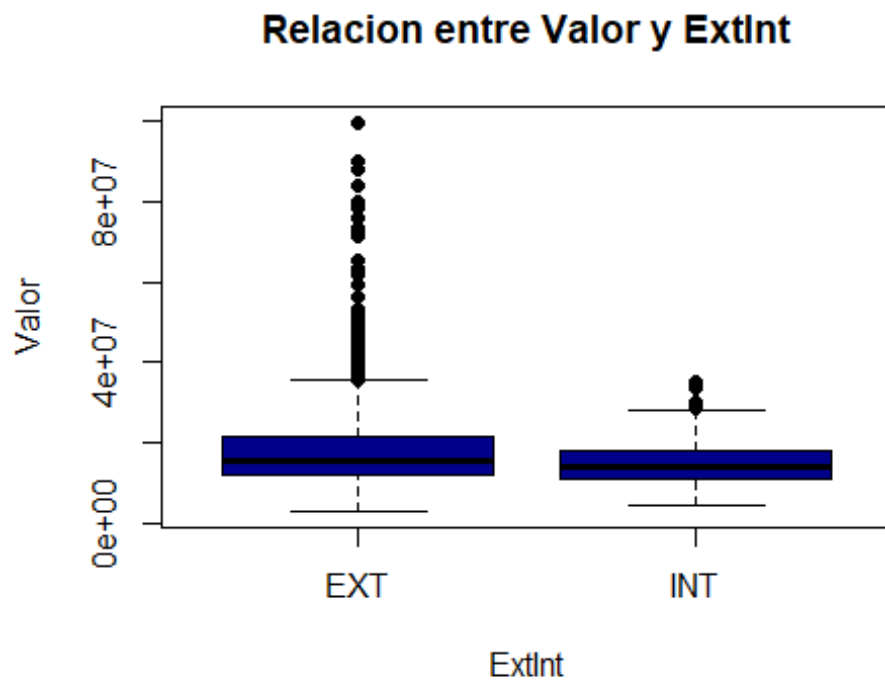
```
# Valor con relación a Tipus  
plot(training_bcnpisos$Tipus, training_bcnpisos$Valor, xlab = "Tipus",  
      ylab = "Valor", main = "Relacion entre Valor y Tipus", type = "p",  
      pch = 19, col = "darkblue")
```



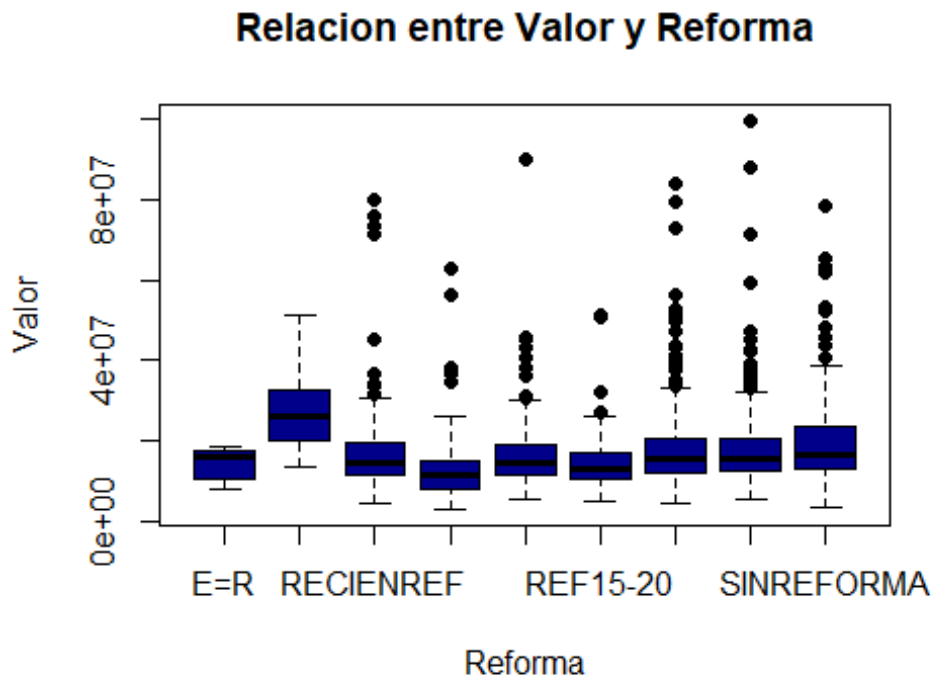
```
# Valor con relación a Ascensor  
plot(training_bcnpisos$Ascens, training_bcnpisos$Valor, xlab = "Ascensor",  
      ylab = "Valor", main = "Relacion entre Valor y Ascensor", type = "p",  
      pch = 19, col = "darkblue")
```



```
# Valor con relación a ExtInt  
plot(training_bcnpisos$ExtInt, training_bcnpisos$Valor, xlab = "ExtInt",  
      ylab = "Valor", main = "Relacion entre Valor y ExtInt", type = "p",  
      pch = 19, col = "darkblue")
```



```
# Valor con relación a Reforma  
plot(training_bcnpisos$Reforma, training_bcnpisos$Valor, xlab = "Reforma",  
ylab = "Valor", main = "Relacion entre Valor y Reforma", type = "p",  
pch = 19, col = "darkblue")
```



```
#Correlación de Valor con Las demás variables numericas de La muestra de
Training
cor(training_bcnpisos$Valor,training_bcnpisos[,unlist(lapply(training_bcn
pisos,is.numeric))])
```

```
##      Valor      Superf      Dorm      Banys      Edat      ValSol
## [1,]      1 0.8803614 0.568745 0.7565376 -0.2268865 0.6330244
```

5. Efectúe la regresión simple de la variable “Valor” respecto de la “Superficie”. A continuación añada a la regresión la variable “Número de dormitorios”. Es significativa esta variable una vez que el modelo ya contiene la variable “Superficie”.

```
#Parte 1: regresión de Valor por La Superficie
regresion_valor <- lm(Valor ~ Superf, data = training_bcnpisos)
summary(regresion_valor)
```

```
##
## Call:
## lm(formula = Valor ~ Superf, data = training_bcnpisos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -40162169  -2337209  -245670   1998325   50428909
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -4195526      305792  -13.72   <2e-16 ***
## Superf      261906       3377    77.56   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4742000 on 1746 degrees of freedom
## Multiple R-squared:  0.775, Adjusted R-squared:  0.7749
## F-statistic: 6015 on 1 and 1746 DF, p-value: < 2.2e-16

#Parte 2: regresión de Valor por La Superficie y Numero de Dormitorios
regresion_valor_dorm <- lm(Valor ~ Superf + Dorm,data =
training_bcnpisos)
summary(regresion_valor_dorm)

##
## Call:
## lm(formula = Valor ~ Superf + Dorm, data = training_bcnpisos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -41720991 -2291526  -275742   1955136  50089376
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3784719     376101  -10.063   <2e-16 ***
## Superf      267521       4513   59.273   <2e-16 ***
## Dorm       -304212     162371   -1.874    0.0612 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4738000 on 1745 degrees of freedom
## Multiple R-squared:  0.7755, Adjusted R-squared:  0.7752
## F-statistic: 3014 on 2 and 1745 DF, p-value: < 2.2e-16
```

Resp.: El resultado nos explica que añadir Dormitorios al modelo no es significativo. La variable Dormitorios es redundante, pues esta muy correlacionada a la variable Superficie. Este problema en estadística es llamado Colinealidad de regresores.

6. Efectúe la regresión múltiple del “Valor” respecto el resto de variables del fichero “BCN_pisos”.¿Le parece que alguna variable predictora es no significativa?

```
regresion_valor_tot <- lm(Valor ~.,data = training_bcnpisos)
summary(regresion_valor_tot)

##
## Call:
## lm(formula = Valor ~ ., data = training_bcnpisos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -16155273 -1161867 91958 1198293 28148468
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.284e+07 1.054e+06 -12.176 < 2e-16 ***
## Superf 2.046e+05 2.939e+03 69.622 < 2e-16 ***
## Dorm -1.572e+05 8.963e+04 -1.754 0.079585 .
## Banyas 1.978e+06 1.624e+05 12.179 < 2e-16 ***
## Edat -2.003e+04 3.300e+03 -6.070 1.57e-09 ***
## Estat2_M 1.271e+06 1.933e+05 6.572 6.56e-11 ***
## Estat3_N 2.118e+06 2.115e+05 10.017 < 2e-16 ***
## Estat4_B 3.313e+06 2.659e+05 12.457 < 2e-16 ***
## Estat5_MB 2.984e+06 4.840e+05 6.165 8.78e-10 ***
## PlantaBajos -4.130e+05 3.203e+05 -1.289 0.197434
## PlantaPlanta -4.440e+05 1.934e+05 -2.296 0.021822 *
## DistEixample -7.749e+05 2.927e+05 -2.647 0.008199 **
## DistGracia -9.975e+05 3.255e+05 -3.064 0.002217 **
## DistHorta -1.092e+06 3.265e+05 -3.346 0.000838 ***
## DistLes_Corts 5.863e+05 4.276e+05 1.371 0.170527
## DistNou_Barris -7.738e+05 3.289e+05 -2.353 0.018753 *
## DistSant_Andreu -9.493e+05 3.272e+05 -2.902 0.003761 **
## DistSant_Marti -1.029e+06 2.923e+05 -3.522 0.000440 ***
## DistSants -9.579e+05 2.947e+05 -3.250 0.001175 **
## DistSarria 3.356e+05 3.886e+05 0.863 0.387987
## ValSol 9.432e+01 2.388e+00 39.498 < 2e-16 ***
## TipusMANZ -1.095e+06 1.927e+05 -5.682 1.56e-08 ***
## AscensSI -1.258e+05 1.528e+05 -0.823 0.410348
## ExtIntINT 4.230e+05 1.747e+05 2.422 0.015552 *
## ReformaOBRANUEVA 1.929e+06 1.086e+06 1.777 0.075787 .
## ReformaRECIENREF 3.289e+05 9.565e+05 0.344 0.730959
## ReformaREF>20 4.137e+05 9.730e+05 0.425 0.670719
## ReformaREF10-15 2.616e+05 9.546e+05 0.274 0.784071
## ReformaREF15-20 3.000e+05 9.788e+05 0.307 0.759241
## ReformaREF1A5 5.876e+05 9.356e+05 0.628 0.530085
## ReformaREF5A10 6.792e+05 9.420e+05 0.721 0.471012
## ReformaSINREFORMA -2.832e+05 9.372e+05 -0.302 0.762532
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2442000 on 1716 degrees of freedom
## Multiple R-squared: 0.9414, Adjusted R-squared: 0.9403
## F-statistic: 888.9 on 31 and 1716 DF, p-value: < 2.2e-16
```

```
anova(regresion_valor_tot)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: Valor
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## Superf 1 1.3525e+17 1.3525e+17 22687.1103 < 2.2e-16 ***
```

```
## Dorm      1 7.8812e+13 7.8812e+13      13.2201 0.0002851 ***
## Banyas    1 6.5636e+15 6.5636e+15    1100.9884 < 2.2e-16 ***
## Edat      1 2.1350e+15 2.1350e+15     358.1239 < 2.2e-16 ***
## Estat     4 2.4615e+15 6.1538e+14     103.2247 < 2.2e-16 ***
## Planta    2 1.3729e+14 6.8647e+13      11.5150 1.077e-05 ***
## Dist      9 8.0693e+15 8.9659e+14     150.3962 < 2.2e-16 ***
## ValSol    1 9.0851e+15 9.0851e+15    1523.9532 < 2.2e-16 ***
## Tipus     1 2.0033e+14 2.0033e+14      33.6036 8.025e-09 ***
## Ascens    1 3.5305e+12 3.5305e+12       0.5922 0.4416710
## ExtInt    1 4.5747e+13 4.5747e+13       7.6738 0.0056633 **
## Reforma   8 2.4778e+14 3.0972e+13       5.1954 1.963e-06 ***
## Residuals 1716 1.0230e+16 5.9615e+12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Resp.: Al principio, nos parece que las variables Dormitorio, Ascensor, y algunas categorías de Reforma y Planta no son significativas. Sin embargo, al ejecutar el Análisis de Variancia, nos parece que solo la variable Ascensor es una variable predictora no significativa.

7. Encuentre la regresión óptima. ¿Cuál es el valor del R2 alcanzado?. ¿Y cuál el valor del R2 por validación cruzada “leave one out”?

#En primer lugar hemos intentado utilizar el paquete olsrr, pero por ser muy time consuming, hemos elegido continuar el ejercicio con el stepwise. Hemos dejado el código con estas instrucciones para efectos de documentación.

```
#library(olsrr)
#bestreg <- ols_step_best_subset(regresion_valor_tot)
```

```
stepwise <- step(regresion_valor_tot)
```

```
## Start: AIC=51451.48
## Valor ~ Superf + Dorm + Banyas + Edat + Estat + Planta + Dist +
##      ValSol + Tipus + Ascens + ExtInt + Reforma
##
##      Df Sum of Sq      RSS   AIC
## - Ascens    1 4.0426e+12 1.0234e+16 51450
## <none>                1.0230e+16 51451
## - Dorm      1 1.8344e+13 1.0248e+16 51453
## - Planta    2 3.1481e+13 1.0261e+16 51453
## - ExtInt    1 3.4961e+13 1.0265e+16 51455
## - Reforma   8 2.4778e+14 1.0478e+16 51477
## - Tipus     1 1.9249e+14 1.0422e+16 51482
## - Dist      9 2.9379e+14 1.0524e+16 51483
## - Edat      1 2.1964e+14 1.0450e+16 51487
## - Banyas    1 8.8423e+14 1.1114e+16 51594
## - Estat     4 1.0198e+15 1.1250e+16 51610
## - ValSol    1 9.3007e+15 1.9531e+16 52580
## - Superf    1 2.8897e+16 3.9127e+16 53794
##
```



```

## Step:  AIC=51450.17
## Valor ~ Superf + Dorm + Banyys + Edat + Estat + Planta + Dist +
##      ValSol + Tipus + ExtInt + Reforma
##
##           Df  Sum of Sq      RSS   AIC
## <none>                1.0234e+16 51450
## - Dorm      1 1.8924e+13 1.0253e+16 51451
## - Planta    2 3.1970e+13 1.0266e+16 51452
## - ExtInt     1 3.3028e+13 1.0267e+16 51454
## - Reforma   8 2.5004e+14 1.0484e+16 51476
## - Tipus     1 1.9027e+14 1.0424e+16 51480
## - Dist      9 2.9317e+14 1.0527e+16 51482
## - Edat      1 2.2879e+14 1.0463e+16 51487
## - Banyys    1 8.8357e+14 1.1118e+16 51593
## - Estat     4 1.0158e+15 1.1250e+16 51608
## - ValSol    1 9.3588e+15 1.9593e+16 52583
## - Superf    1 2.9590e+16 3.9824e+16 53823

stepwise

##
## Call:
## lm(formula = Valor ~ Superf + Dorm + Banyys + Edat + Estat + Planta +
##      Dist + ValSol + Tipus + ExtInt + Reforma, data =
training_bcnpisos)
##
## Coefficients:
##      (Intercept)          Superf          Dorm
Banyys
##      -1.288e+07      2.042e+05     -1.596e+05
1.977e+06
##      Edat      Estat2_M      Estat3_N
Estat4_B
##      -1.904e+04      1.264e+06      2.105e+06
3.302e+06
##      Estat5_MB      PlantaBajos      PlantaPlanta
DistEixample
##      2.957e+06     -4.005e+05     -4.478e+05      -
7.765e+05
##      DistGracia      DistHorta      DistLes_Corts
DistNou_Barris
##      -9.772e+05     -1.049e+06      6.041e+05      -
7.412e+05
##      DistSant_Andreu      DistSant_Marti      DistSants
DistSarria
##      -9.356e+05     -1.019e+06     -9.533e+05
3.542e+05
##      ValSol      TipusMANZ      ExtIntINT
ReformaOBRANUEVA
##      9.412e+01     -1.088e+06      4.093e+05

```

```

1.935e+06
## ReformaRECIENTREF      ReformaREF>20      ReformaREF10-15
ReformaREF15-20
##          3.197e+05          3.904e+05          2.441e+05
2.804e+05
## ReformaREF1A5      ReformaREF5A10      ReformaSINREFORMA
##          5.734e+05          6.548e+05          -3.042e+05

```

summary(stepwise)

```

##
## Call:
## lm(formula = Valor ~ Superf + Dorm + Banys + Edat + Estat + Planta +
##      Dist + ValSol + Tipus + ExtInt + Reforma, data =
training_bcnpisos)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -16160138  -1169276    97788   1194856  28196036
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.288e+07  1.053e+06 -12.226 < 2e-16 ***
## Superf         2.042e+05  2.898e+03  70.459 < 2e-16 ***
## Dorm          -1.596e+05  8.957e+04  -1.782 0.074952 .
## Banys          1.977e+06  1.624e+05  12.175 < 2e-16 ***
## Edat          -1.904e+04  3.073e+03  -6.196 7.25e-10 ***
## Estat2_M       1.264e+06  1.932e+05   6.545 7.85e-11 ***
## Estat3_N       2.105e+06  2.108e+05   9.984 < 2e-16 ***
## Estat4_B       3.302e+06  2.656e+05  12.433 < 2e-16 ***
## Estat5_MB      2.957e+06  4.828e+05   6.124 1.13e-09 ***
## PlantaBajos   -4.005e+05  3.199e+05  -1.252 0.210866
## PlantaPlanta  -4.478e+05  1.934e+05  -2.316 0.020699 *
## DistEixample  -7.765e+05  2.927e+05  -2.653 0.008059 **
## DistGracia    -9.772e+05  3.246e+05  -3.011 0.002644 **
## DistHorta     -1.049e+06  3.223e+05  -3.256 0.001151 **
## DistLes_Corts  6.041e+05  4.270e+05   1.415 0.157288
## DistNou_Barris -7.412e+05  3.265e+05  -2.270 0.023317 *
## DistSant_Andreu -9.356e+05  3.267e+05  -2.864 0.004239 **
## DistSant_Marti -1.019e+06  2.920e+05  -3.490 0.000495 ***
## DistSants     -9.533e+05  2.946e+05  -3.236 0.001237 **
## DistSarria     3.542e+05  3.879e+05   0.913 0.361313
## ValSol         9.412e+01  2.375e+00  39.625 < 2e-16 ***
## TipusMANZ     -1.088e+06  1.925e+05  -5.650 1.88e-08 ***
## ExtIntINT      4.093e+05  1.739e+05   2.354 0.018685 *
## ReformaOBRANUEVA 1.935e+06  1.085e+06   1.782 0.074887 .
## ReformaRECIENTREF 3.197e+05  9.563e+05   0.334 0.738218
## ReformaREF>20   3.904e+05  9.724e+05   0.401 0.688126
## ReformaREF10-15 2.441e+05  9.543e+05   0.256 0.798100
## ReformaREF15-20 2.804e+05  9.784e+05   0.287 0.774443

```

```
## ReformaREF1A5      5.734e+05  9.354e+05   0.613 0.539936
## ReformaREF5A10     6.548e+05  9.415e+05   0.696 0.486810
## ReformaSINREFORMA -3.042e+05  9.367e+05  -0.325 0.745403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2441000 on 1717 degrees of freedom
## Multiple R-squared:  0.9414, Adjusted R-squared:  0.9403
## F-statistic: 918.7 on 30 and 1717 DF,  p-value: < 2.2e-16

regresion_valor_tot_step <- lm(formula = Valor ~ Superf + Banys + Edat +
Estat + Planta + Dist + ValSol + Tipus + ExtInt + Reforma, data =
training_bcnpisos)

PRESS <- sum((regresion_valor_tot_step$residuals/(1-
ls.diag(regresion_valor_tot_step)$hat))^2)
R2loo <- 1-PRESS/(var(training_bcnpisos$Valor)*(nrow(training_bcnpisos)-
1))
R2loo

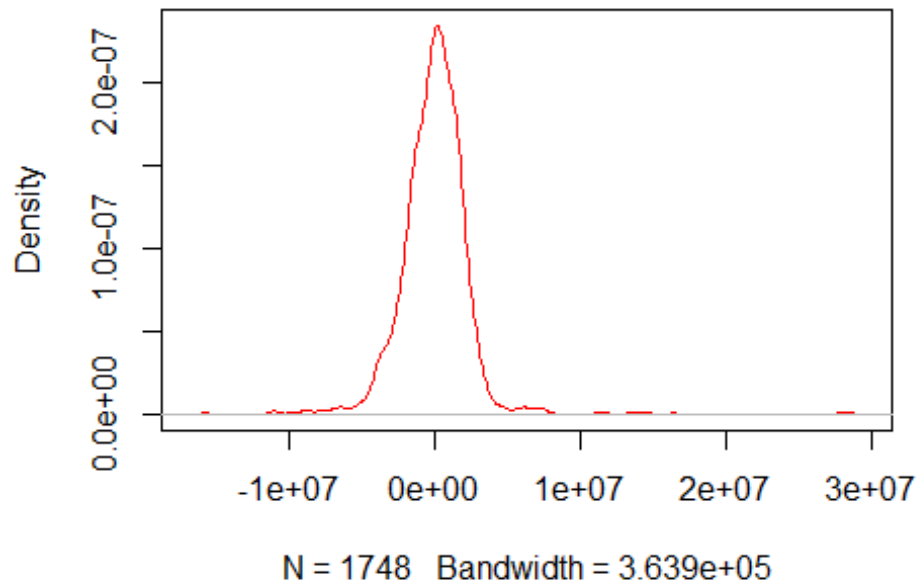
## [1] 0.9371742
```

Resp.: El valor de R2 alcanzado es de 94,59%, y el valor alcanzado para R2 Leave One Out es de 94,18%

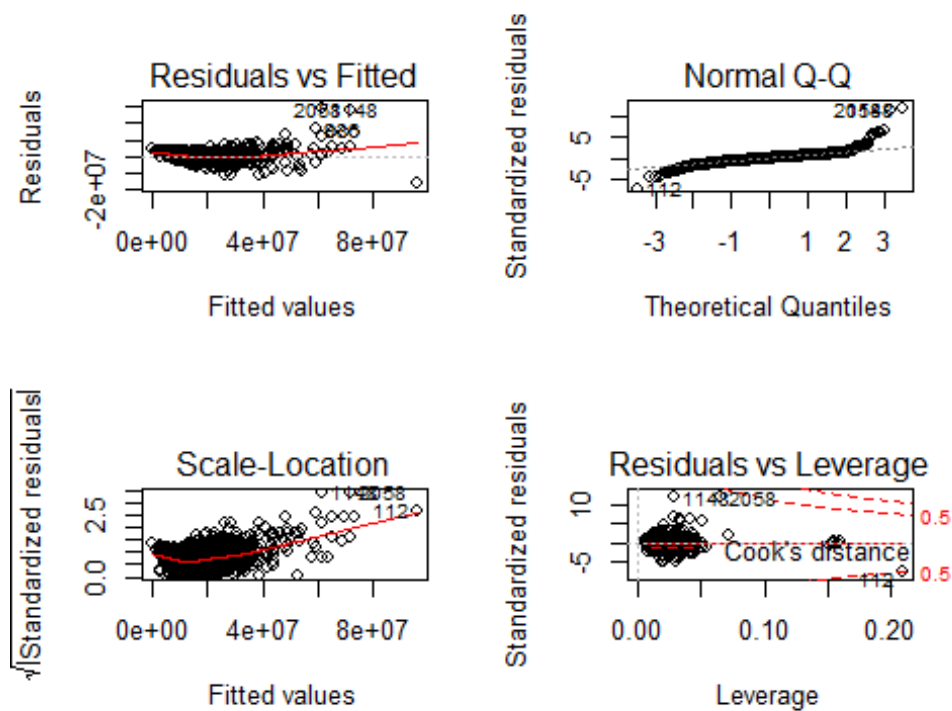
8. Realice el análisis de los residuos. ¿Son normales los residuos?, ¿Existe alguna relación de dependencia con los valores ajustados?. ¿Existe heterocedasticidad?. ¿Existen observaciones influyentes?

```
plot(density(regresion_valor_tot_step$residuals),col="red")
```

```
density.default(x = regresion_valor_tot_step$residu
```



```
par(mfrow = c(2, 2))
plot(regresion_valor_tot_step)
```



```
par(mfrow = c(1, 1))
```

Para contestar las preguntas: 1) ¿Son normales los residuos? -> Resp.: A través del análisis de los residuos, vemos que son normales. 2) ¿Existe alguna relación de dependencia con los valores ajustados? -> Resp.: Si, es notable en el Scale-Location. 3) ¿Existe heterocedasticidad? -> Resp.: Si, existe. A través del Scale-Location se puede ver que hay valores variados en el modelo. 4) ¿Existen observaciones influyentes? -> Resp.: Mirando el Residuals vs Leverage (Cooks Distance) vemos observaciones que son potenciales outliers, sin embargo que no son observaciones influyentes.

9. Obtenga el valor del R2 de predicción en la muestra test.

```
pred <- predict(regresion_valor_tot_step, newdata = test_bcnpisos)
predSSE <- sum((test_bcnpisos$Valor-pred)^2)
R2Test <- 1-(predSSE)/(var(test_bcnpisos$Valor)*(nrow(test_bcnpisos)-1))
R2Test

## [1] 0.9496189
```

10. Obtenga el fichero con las predicciones del valor de las viviendas con su intervalo de confianza del 95%, para los pisos de la muestra test.

```
pred_int <- predict(regresion_valor_tot_step, newdata = test_bcnpisos,
interval = "confidence")
head(pred_int)
```

##		fit	lwr	upr
## 4		8047797	7428482	8667112
## 6		8241642	7513214	8970071
## 17		13387606	12699169	14076044
## 20		6282076	5621773	6942379
## 31		14251192	13693616	14808768
## 38		5507523	4777220	6237826

Resp.: No hemos podido encontrar predicciones del valor de las viviendas con su intervalo de confianza del 95% para los pisos de la muestra test. Los valores resultantes eran muy parecidos con la muestra de valores que hemos imprimido para completar este ejercicio.