FIFAS

FIFAS

FIFAS

```
#cargo librerias

# cargamos datos
dat <- read.csv("FIFA_WC_Russia_DB.csv" , header = TRUE)
```

```
summary(dat)
```

```
 Observation      Country          Mkt_Value_per_player   Mkt_Value
Min.   : 1.00   Length:32          Length:32            Min.   :    7.35
1st Qu.: 8.75   Class :character   Class :character     1st Qu.:   77.37
Median :16.50   Mode  :character   Mode  :character     Median :  257.57
Mean   :16.50                                           Mean   :  681.49
3rd Qu.:24.25                                           3rd Qu.:  383.07
Max.   :32.00                                           Max.   :12300.00
  Player_Age    Player_Height     Coach_Wage          Bets
Min.   :25.90   Min.   :176.2   Min.   :0.2000   Min.   : 0.000
1st Qu.:27.20   1st Qu.:180.2   1st Qu.:0.4875   1st Qu.: 0.000
Median :28.10   Median :181.8   Median :1.0250   Median : 0.000
Mean   :27.86   Mean   :181.9   Mean   :1.3513   Mean   : 2.344
3rd Qu.:28.60   3rd Qu.:183.8   3rd Qu.:1.8500   3rd Qu.: 0.500
Max.   :29.60   Max.   :185.6   Max.   :3.8500   Max.   :23.000
  Goals_For      Goals_Against    Invested_Money  Squad_Formation
Min.   : 2.000   Min.   : 2.000   Min.   :12.11   Min.   :0.0000
1st Qu.: 2.000   1st Qu.: 4.000   1st Qu.:32.23   1st Qu.:0.0000
Median : 3.500   Median : 5.000   Median :42.18   Median :0.0000
Mean   : 5.281   Mean   : 5.281   Mean   :40.63   Mean   :0.4688
3rd Qu.: 6.250   3rd Qu.: 6.250   3rd Qu.:49.84   3rd Qu.:1.0000
Max.   :16.000   Max.   :11.000   Max.   :62.40   Max.   :1.0000
```

```
   Games_Won
 Min.   :0.000
 1st Qu.:1.000
 Median :1.000
 Mean   :1.688
 3rd Qu.:2.000
 Max.   :7.000
```

```r
dat <- read.csv("FIFA_WC_Russia_DB.csv" , header = TRUE)

#Limpieza de Datos
typeof(dat$Country) #character
```

```
[1] "character"
```

```r
dat$Country <- as.factor(dat$Country)
typeof(dat$Country)
```

```
[1] "integer"
```

```r
dat$Mkt_Value_per_player <- as.numeric(dat$Mkt_Value_per_player)
```

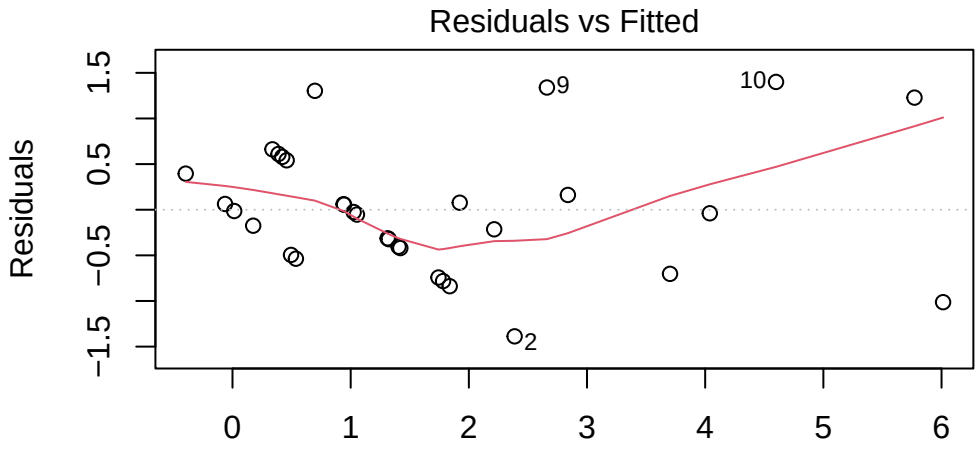
```
Warning: NAs introduced by coercion
```

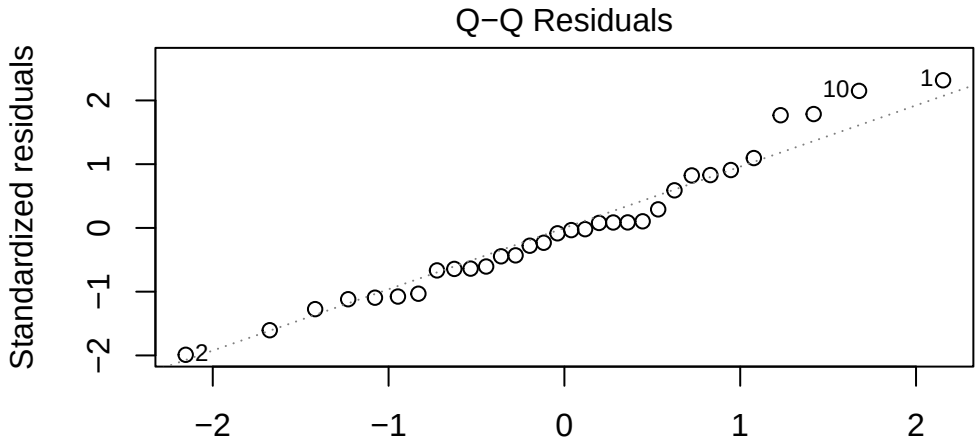
```r
dat$Mkt_Value_per_player[1] <- 1080
dat$Mkt_Value_per_player[2] <- 1040

dat$Squad_Formation <- as.factor(dat$Squad_Formation)

#Modelos de Regresion
#numero de victorias
modelo1 <- lm(Games_Won ~ Mkt_Value_per_player + Mkt_Value + Bets + Player_Age + Player_He
summary(modelo1)
```

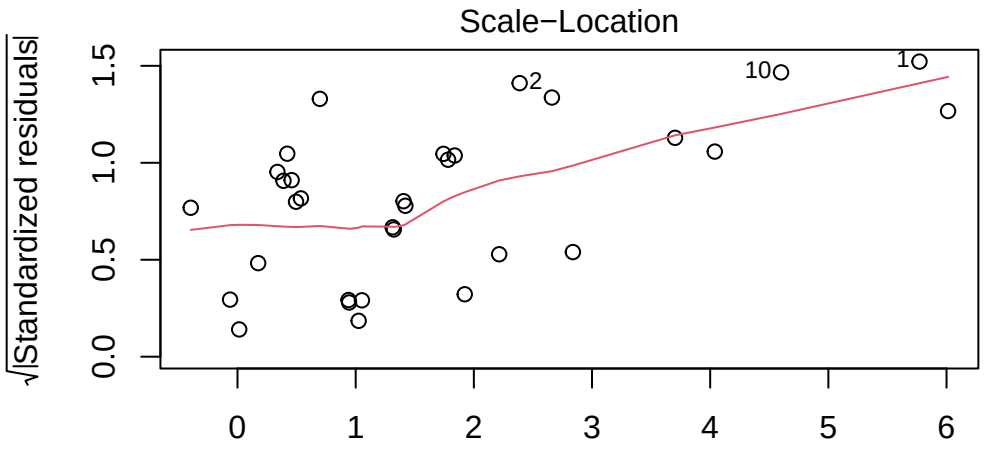
```
Call:
lm(formula = Games_Won ~ Mkt_Value_per_player + Mkt_Value + Bets +
```

```
    Player_Age + Player_Height + Coach_Wage + Goals_For + Goals_Against +
    Invested_Money + Squad_Formation, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-1.38742 -0.43896 -0.03197  0.43233  1.40050

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)           1.503e+01  1.775e+01   0.847    0.407
Mkt_Value_per_player  1.723e-04  1.193e-03   0.144    0.887
Mkt_Value            -5.749e-05  8.956e-05  -0.642    0.528
Bets                 -1.952e-02  5.905e-02  -0.330    0.744
Player_Age           -2.181e-01  1.982e-01  -1.100    0.284
Player_Height        -4.842e-02  8.298e-02  -0.584    0.566
Coach_Wage            4.720e-02  3.007e-01   0.157    0.877
Goals_For             4.017e-01  6.498e-02   6.181 3.93e-06 ***
Goals_Against        -9.190e-02  8.541e-02  -1.076    0.294
Invested_Money       -5.761e-03  1.663e-02  -0.346    0.733
Squad_Formation1      2.232e-01  3.592e-01   0.621    0.541
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8471 on 21 degrees of freedom
Multiple R-squared:  0.8444,    Adjusted R-squared:  0.7704
F-statistic:  11.4 on 10 and 21 DF,  p-value: 2.083e-06
```

```r
#standardCoefs(modelo1)
plot(modelo1)
```

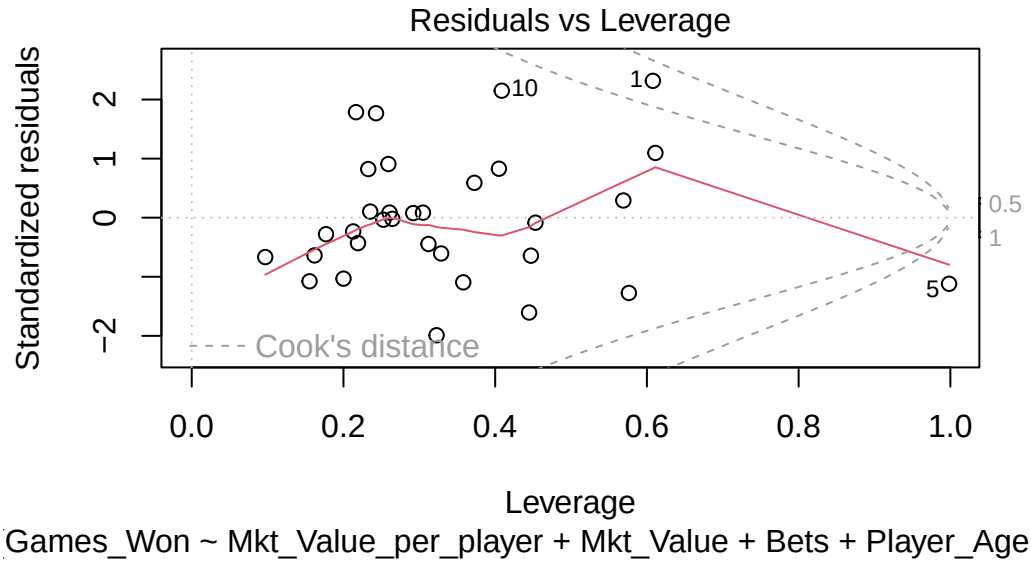

Residuals vs Fitted
Residuals
1.5
0.5
−0.5
−1.5
9
10
2
0
1
2
3
4
5
6
Fitted values
Games_Won ~ Mkt_Value_per_player + Mkt_Value + Bets + Player_Age +


Q−Q Residuals
Standardized residuals
2
1
0
−1
−2
10
1
2
−2
−1
0
1
2
Theoretical Quantiles
Games_Won ~ Mkt_Value_per_player + Mkt_Value + Bets + Player_Age +




```
Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

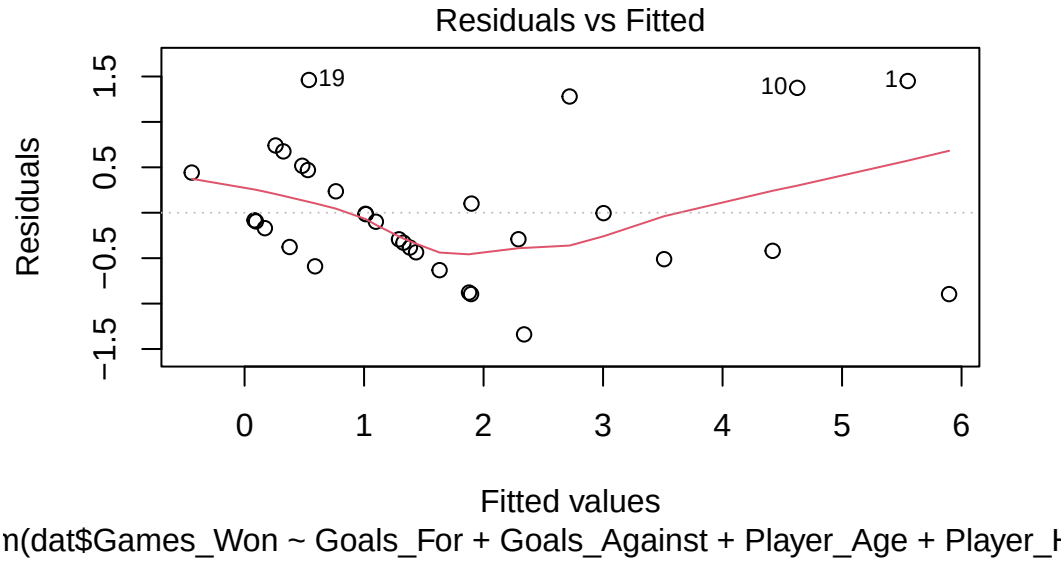
```
Warning in sqrt(crit * p * (1 - hh)/hh): NaNs produced
```

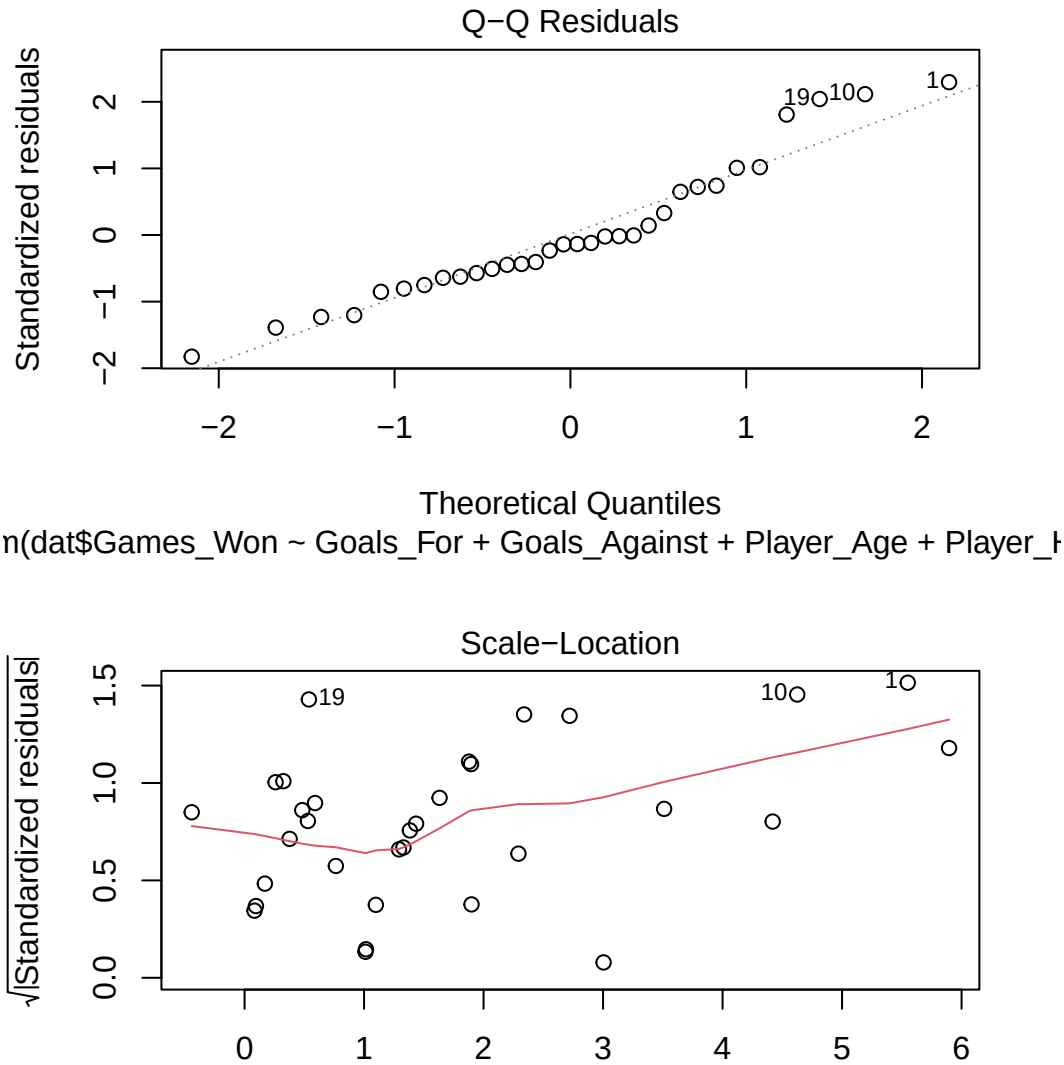



```
modelo2 <- lm(dat$Games_Won ~ Goals_For + Goals_Against + Player_Age + Player_Height, data
summary(modelo2)
```

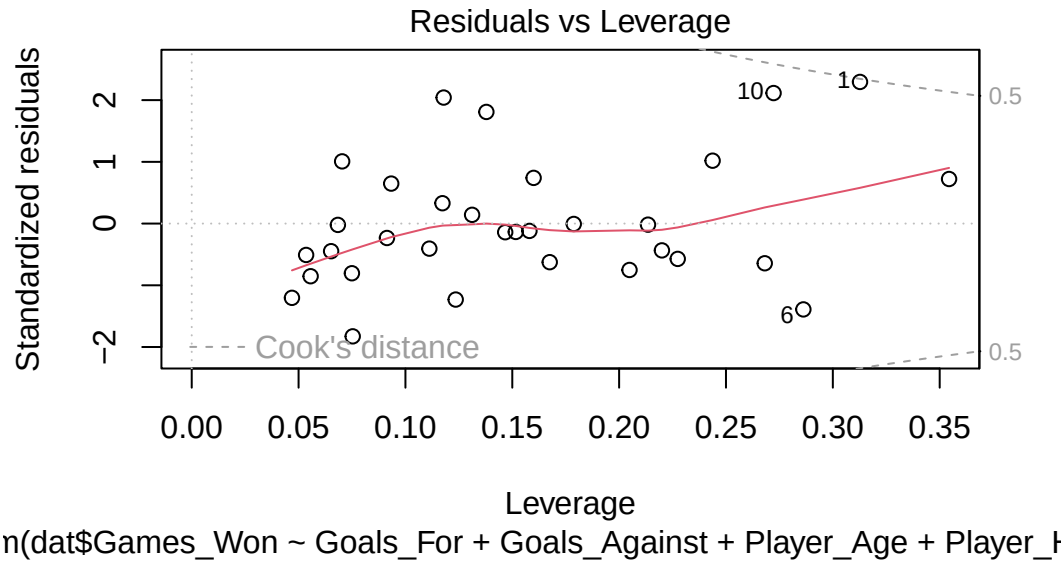
```
Call:
lm(formula = dat$Games_Won ~ Goals_For + Goals_Against + Player_Age +
    Player_Height, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-1.33920 -0.42330 -0.09696  0.44922  1.46158

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   15.18335   13.14455   1.155    0.258
Goals_For      0.40599    0.03762  10.792  2.7e-11 ***
Goals_Against -0.11483    0.07038  -1.632    0.114
Player_Age    -0.17681    0.15624  -1.132    0.268
Player_Height -0.05557    0.06148  -0.904    0.374
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7623 on 27 degrees of freedom
Multiple R-squared:  0.8381,    Adjusted R-squared:  0.8141
F-statistic: 34.93 on 4 and 27 DF,  p-value: 2.611e-10
```

```r
#standardCoefs(modelo2)
plot(modelo2)
```




Q−Q Residuals
Standardized residuals
−2
−1
0
1
2
19
10
1
Theoretical Quantiles
n(dat$Games_Won ~ Goals_For + Goals_Against + Player_Age + Player_H

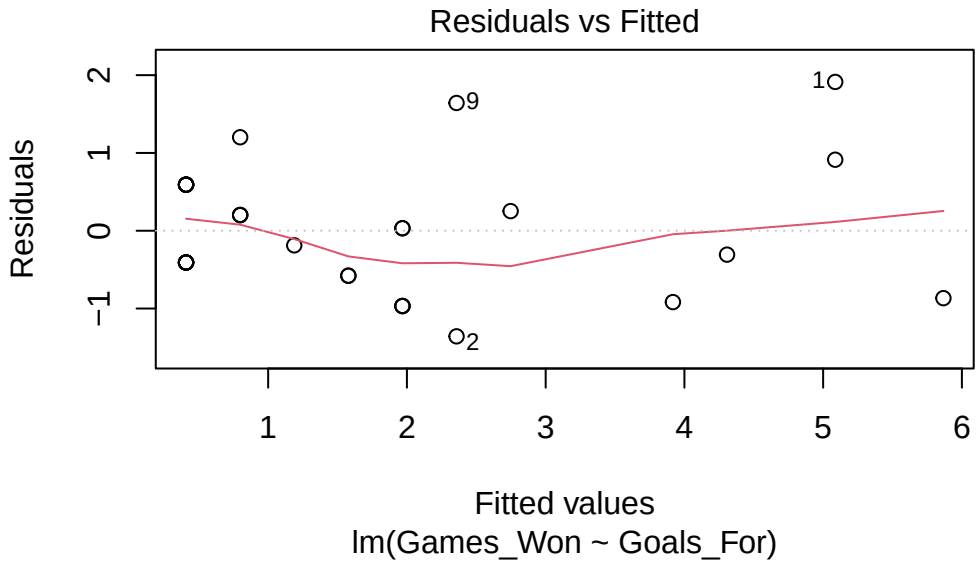
Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
0
1
2
3
4
5
6
19
10
1
Fitted values
n(dat$Games_Won ~ Goals_For + Goals_Against + Player_Age + Player_H

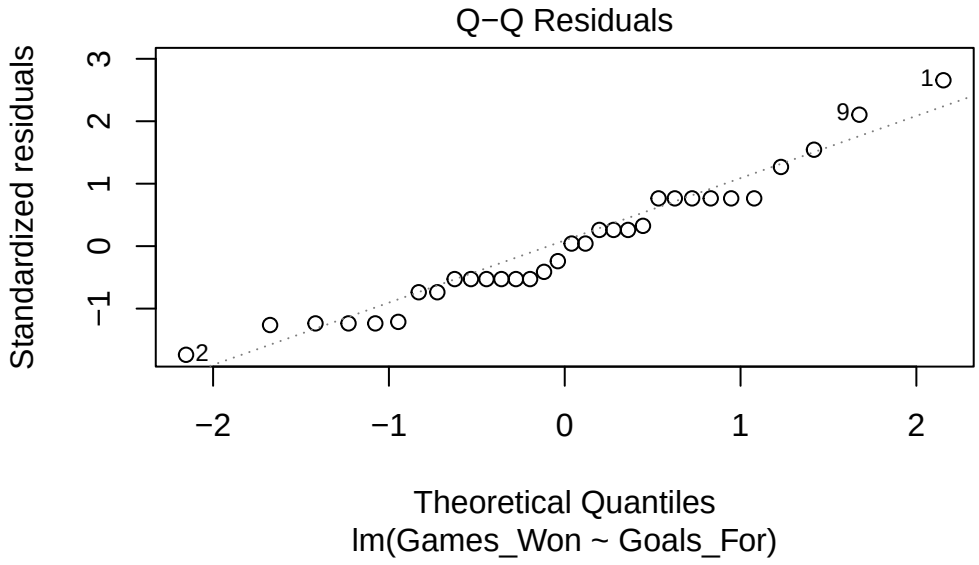


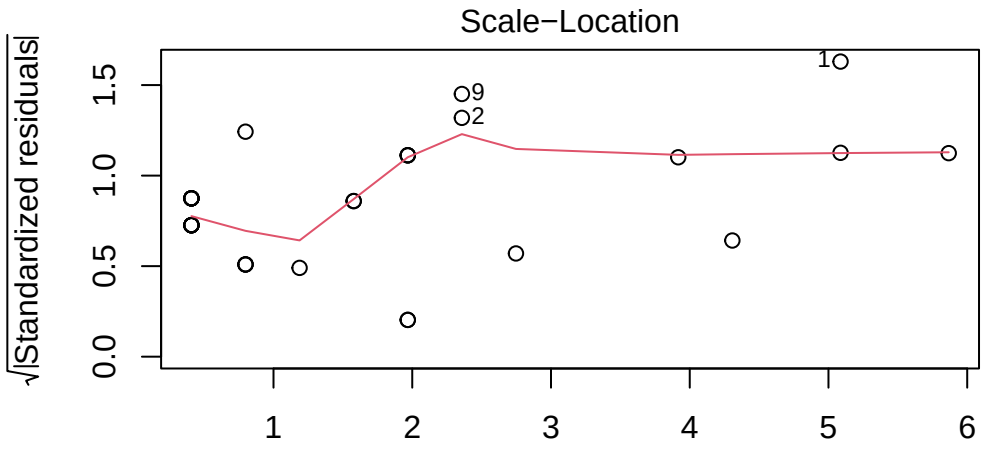

```r
modelo3 <- lm(Games_Won ~ Goals_For,data = dat)
summary(modelo3)
```

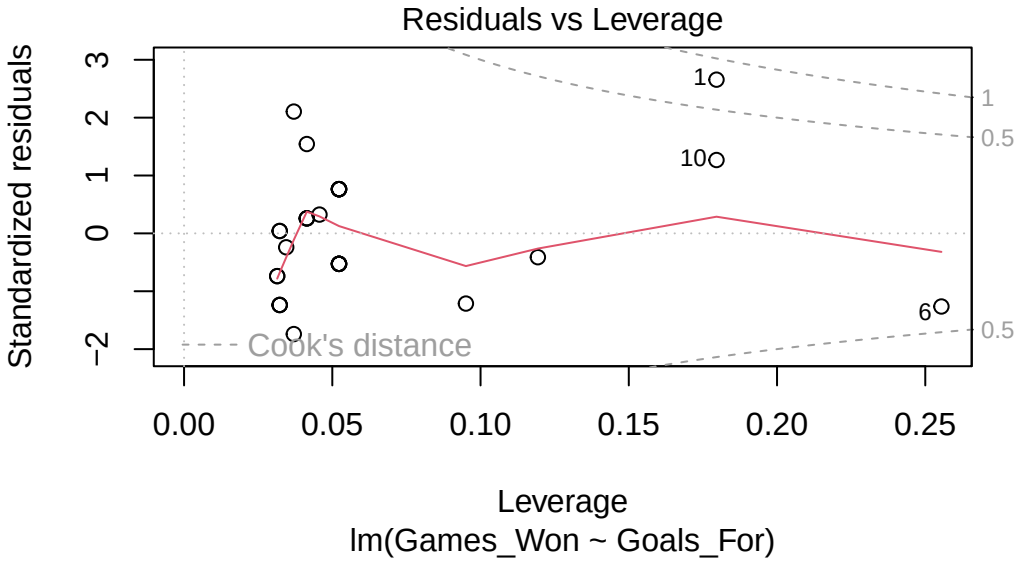
```
Call:
lm(formula = Games_Won ~ Goals_For, data = dat)

Residuals:
     Min       1Q   Median       3Q      Max
-1.35764 -0.45056 -0.07784  0.59187  1.91304

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.37167    0.23274  -1.597    0.121
Goals_For    0.38990    0.03512  11.100 3.81e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7951 on 30 degrees of freedom
Multiple R-squared:  0.8042,    Adjusted R-squared:  0.7977
F-statistic: 123.2 on 1 and 30 DF,  p-value: 3.81e-12
```

```r
plot(modelo3)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Games_Won ~ Goals_For)


Q−Q Residuals
Standardized residuals
Theoretical Quantiles
lm(Games_Won ~ Goals_For)






```r
# Usar step() para encontrar el mejor modelo
mejor_modeloAIC <- step(modelo1, direction = "both")  # Puede ser "forward", "backward", o
```

```
Start:  AIC=-2.1
Games_Won ~ Mkt_Value_per_player + Mkt_Value + Bets + Player_Age +
    Player_Height + Coach_Wage + Goals_For + Goals_Against +
    Invested_Money + Squad_Formation
```

```
                       Df Sum of Sq    RSS     AIC
- Mkt_Value_per_player  1    0.0150 15.085 -4.0642
- Coach_Wage            1    0.0177 15.088 -4.0584
- Bets                  1    0.0784 15.149 -3.9299
- Invested_Money        1    0.0861 15.157 -3.9136
- Player_Height         1    0.2444 15.315 -3.5812
- Squad_Formation       1    0.2770 15.348 -3.5130
- Mkt_Value             1    0.2957 15.366 -3.4741
- Goals_Against         1    0.8308 15.901 -2.3788
- Player_Age            1    0.8688 15.939 -2.3023
<none>                             15.070 -2.0959
- Goals_For             1   27.4209 42.491 29.0741

Step:  AIC=-4.06
Games_Won ~ Mkt_Value + Bets + Player_Age + Player_Height + Coach_Wage +
    Goals_For + Goals_Against + Invested_Money + Squad_Formation

                       Df Sum of Sq    RSS    AIC
- Coach_Wage            1     0.022 15.108 -6.017
- Bets                  1     0.068 15.154 -5.919
- Invested_Money        1     0.071 15.157 -5.913
- Squad_Formation       1     0.264 15.349 -5.510
- Mkt_Value             1     0.282 15.367 -5.472
- Player_Height         1     0.295 15.380 -5.445
- Goals_Against         1     0.857 15.943 -4.296
<none>                             15.085 -4.064
- Player_Age            1     0.974 16.060 -4.061
+ Mkt_Value_per_player  1     0.015 15.070 -2.096
- Goals_For             1    43.451 58.537 37.325

Step:  AIC=-6.02
Games_Won ~ Mkt_Value + Bets + Player_Age + Player_Height + Goals_For +
    Goals_Against + Invested_Money + Squad_Formation

                       Df Sum of Sq    RSS    AIC
- Bets                  1     0.053 15.161 -7.904
- Invested_Money        1     0.072 15.180 -7.864
- Mkt_Value             1     0.279 15.387 -7.431
- Squad_Formation       1     0.320 15.427 -7.347
- Player_Height         1     0.472 15.579 -7.033
- Goals_Against         1     0.963 16.071 -6.040
<none>                             15.108 -6.017
- Player_Age            1     1.003 16.110 -5.960
```

```
+ Coach_Wage            1     0.022 15.085 -4.064
+ Mkt_Value_per_player  1     0.020 15.088 -4.058
- Goals_For             1    54.440 69.548 40.841

Step:  AIC=-7.9
Games_Won ~ Mkt_Value + Player_Age + Player_Height + Goals_For +
    Goals_Against + Invested_Money + Squad_Formation

                       Df Sum of Sq    RSS    AIC
- Invested_Money        1     0.101 15.262 -9.691
- Squad_Formation       1     0.279 15.440 -9.320
- Mkt_Value             1     0.300 15.461 -9.278
- Player_Height         1     0.440 15.601 -8.989
- Goals_Against         1     0.925 16.086 -8.009
<none>                              15.161 -7.904
- Player_Age            1     1.042 16.203 -7.778
+ Bets                  1     0.053 15.108 -6.017
+ Mkt_Value_per_player  1     0.012 15.149 -5.930
+ Coach_Wage            1     0.007 15.154 -5.919
- Goals_For             1    54.475 69.636 38.881

Step:  AIC=-9.69
Games_Won ~ Mkt_Value + Player_Age + Player_Height + Goals_For +
    Goals_Against + Squad_Formation

                       Df Sum of Sq    RSS     AIC
- Squad_Formation       1     0.206 15.468 -11.262
- Mkt_Value             1     0.304 15.567 -11.060
- Player_Height         1     0.570 15.832 -10.518
- Goals_Against         1     0.940 16.202  -9.779
<none>                              15.262  -9.691
- Player_Age            1     1.061 16.323  -9.540
+ Invested_Money        1     0.101 15.161  -7.904
+ Bets                  1     0.082 15.180  -7.864
+ Mkt_Value_per_player  1     0.049 15.214  -7.793
+ Coach_Wage            1     0.016 15.246  -7.726
- Goals_For             1    60.295 75.557  39.493

Step:  AIC=-11.26
Games_Won ~ Mkt_Value + Player_Age + Player_Height + Goals_For +
    Goals_Against

                       Df Sum of Sq    RSS     AIC
```

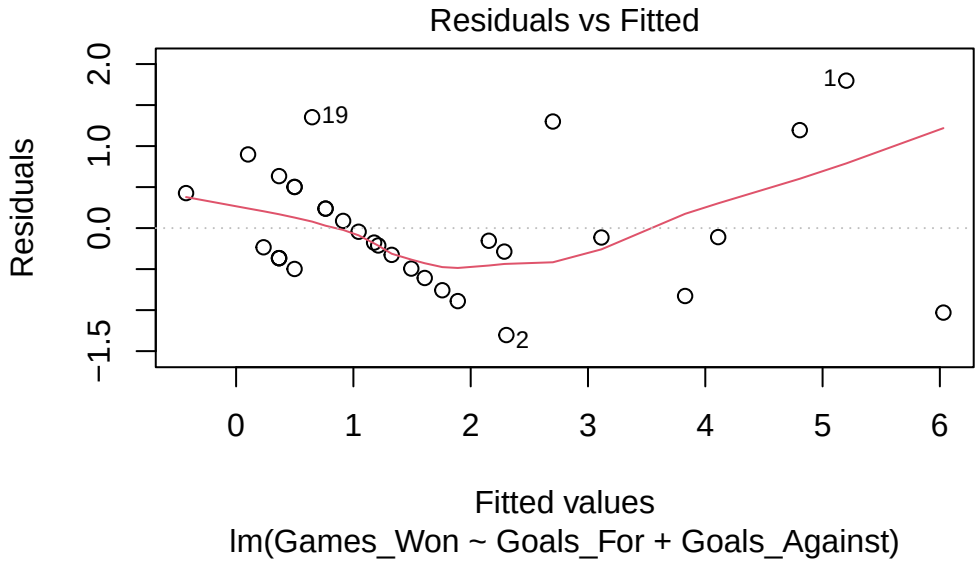
```
- Mkt_Value             1     0.220 15.688 -12.810
- Player_Height         1     0.561 16.030 -12.121
- Player_Age            1     0.958 16.426 -11.339
<none>                              15.468 -11.262
- Goals_Against         1     1.199 16.668 -10.873
+ Squad_Formation       1     0.206 15.262  -9.691
+ Invested_Money        1     0.028 15.440  -9.320
+ Bets                  1     0.023 15.445  -9.310
+ Mkt_Value_per_player  1     0.015 15.454  -9.292
+ Coach_Wage            1     0.002 15.467  -9.266
- Goals_For             1    66.214 81.683  39.987

Step:  AIC=-12.81
Games_Won ~ Player_Age + Player_Height + Goals_For + Goals_Against

                       Df Sum of Sq    RSS     AIC
- Player_Height         1     0.475 16.163 -13.856
- Player_Age            1     0.744 16.433 -13.327
<none>                              15.688 -12.810
- Goals_Against         1     1.547 17.235 -11.801
+ Mkt_Value             1     0.220 15.468 -11.262
+ Squad_Formation       1     0.122 15.567 -11.060
+ Mkt_Value_per_player  1     0.056 15.633 -10.924
+ Bets                  1     0.045 15.643 -10.902
+ Invested_Money        1     0.041 15.647 -10.894
+ Coach_Wage            1     0.001 15.688 -10.811
- Goals_For             1    67.672 83.360  38.638

Step:  AIC=-13.86
Games_Won ~ Player_Age + Goals_For + Goals_Against

                       Df Sum of Sq    RSS     AIC
- Player_Age            1     0.475 16.638 -14.930
<none>                              16.163 -13.856
- Goals_Against         1     1.439 17.602 -13.128
+ Player_Height         1     0.475 15.688 -12.810
+ Mkt_Value             1     0.133 16.030 -12.121
+ Squad_Formation       1     0.130 16.033 -12.115
+ Invested_Money        1     0.118 16.045 -12.091
+ Coach_Wage            1     0.052 16.111 -11.959
+ Bets                  1     0.022 16.142 -11.899
+ Mkt_Value_per_player  1     0.006 16.158 -11.867
- Goals_For             1    67.215 83.378  36.645
```

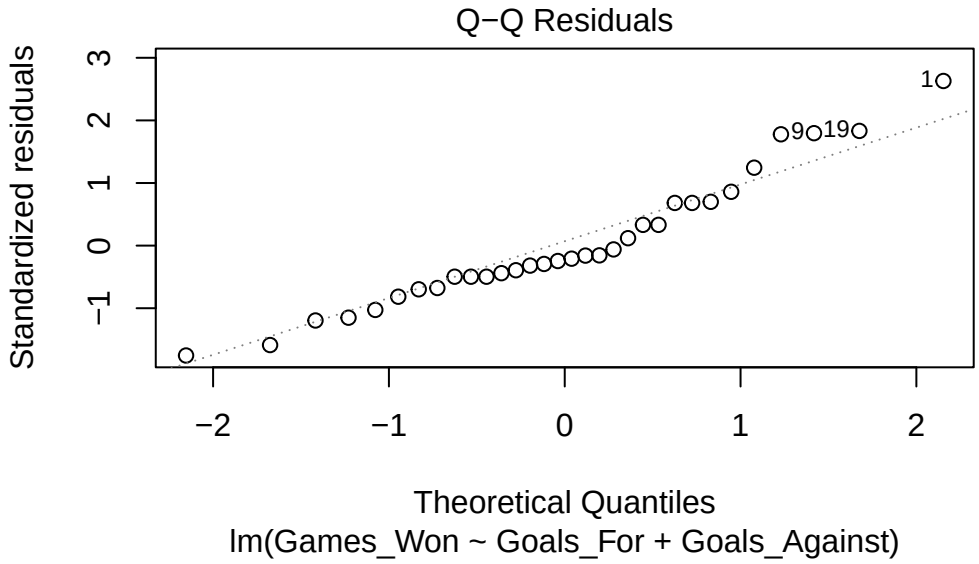
```
Step:  AIC=-14.93
Games_Won ~ Goals_For + Goals_Against

                      Df Sum of Sq    RSS     AIC
<none>                              16.638 -14.930
+ Player_Age           1     0.475 16.163 -13.856
+ Player_Height        1     0.205 16.433 -13.327
+ Invested_Money       1     0.107 16.531 -13.136
+ Squad_Formation      1     0.102 16.536 -13.126
+ Bets                 1     0.044 16.594 -13.015
+ Coach_Wage           1     0.022 16.616 -12.972
+ Mkt_Value            1     0.005 16.633 -12.940
+ Mkt_Value_per_player  1     0.000 16.638 -12.930
- Goals_Against        1     2.330 18.968 -12.736
- Goals_For            1    77.885 94.523  38.659
```

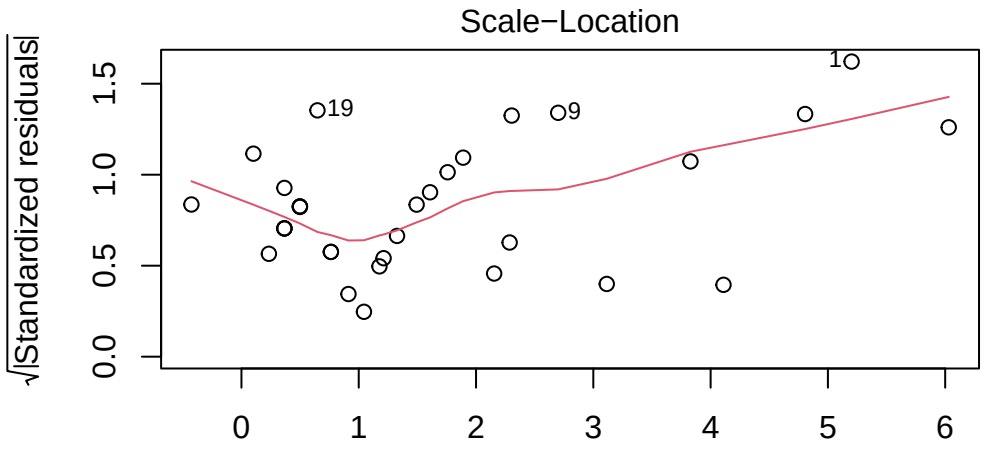
```r
# Ver el resumen del mejor modelo
summary(mejor_modeloAIC)
```

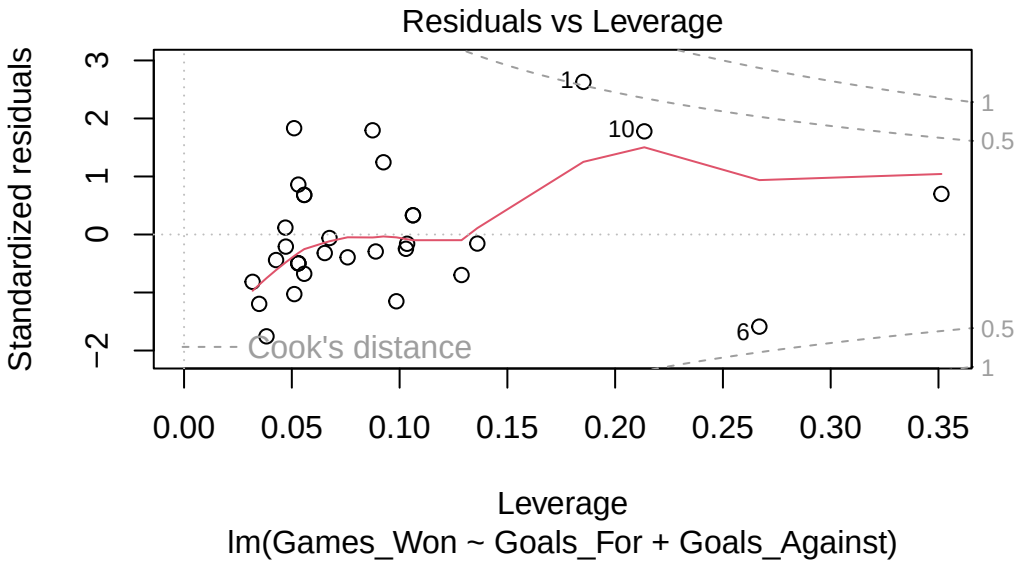
```
Call:
lm(formula = Games_Won ~ Goals_For + Goals_Against, data = dat)

Residuals:
    Min      1Q  Median      3Q     Max
-1.3041 -0.3981 -0.1655  0.4454  1.7979

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.19894    0.35962   0.553   0.5844
Goals_For      0.41400    0.03553  11.651 1.85e-12 ***
Goals_Against -0.13215    0.06557  -2.015   0.0532 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7574 on 29 degrees of freedom
Multiple R-squared:  0.8283,    Adjusted R-squared:  0.8164
F-statistic: 69.93 on 2 and 29 DF,  p-value: 8.051e-12
```

```
plot(mejor_modeloAIC)
```










```r
# Hacer predicciones
predicciones <- predict(mejor_modeloAIC, dat[1,])

# Ver las predicciones
print(predicciones)
```

```
       1
5.202116
```