

Regressão Múltipla

Load data

```
boston = readxl::read_excel("boston.xlsx")
b = boston[, -1]
```

Adicionando labels

```
b$chas=as.factor(b$chas)
levels(b$chas)=c("otherwise", "bounds river")
```

Sumario dos dados

```
summary(b)
```

```
##          crim              zn          indus          chas
## Min.   : 0.00632   Min.   : 0.00   Min.   : 0.46   otherwise :471
## 1st Qu.: 0.08204   1st Qu.: 0.00   1st Qu.: 5.19   bounds river: 35
## Median : 0.25651   Median : 0.00   Median : 9.69
## Mean   : 3.61352   Mean    : 11.36   Mean    :11.14
## 3rd Qu.: 3.67708   3rd Qu.: 12.50   3rd Qu.:18.10
## Max.   :88.97620   Max.    :100.00   Max.    :27.74
##          nox          rm          age          dis
## Min.   :0.3850   Min.   :3.561   Min.   : 2.90   Min.   : 1.130
## 1st Qu.:0.4490   1st Qu.:5.886   1st Qu.: 45.02   1st Qu.: 2.100
## Median :0.5380   Median :6.208   Median : 77.50   Median : 3.207
## Mean   :0.5547   Mean    :6.285   Mean    : 68.57   Mean    : 3.795
## 3rd Qu.:0.6240   3rd Qu.:6.623   3rd Qu.: 94.08   3rd Qu.: 5.188
## Max.   :0.8710   Max.    :8.780   Max.    :100.00   Max.    :12.127
##          rad          tax          ptratio          lstat
## Min.   : 1.000   Min.   :187.0   Min.   :12.60   Min.   : 1.73
## 1st Qu.: 4.000   1st Qu.:279.0   1st Qu.:17.40   1st Qu.: 6.95
## Median : 5.000   Median :330.0   Median :19.05   Median :11.36
## Mean   : 9.549   Mean    :408.2   Mean    :18.46   Mean    :12.65
## 3rd Qu.:24.000   3rd Qu.:666.0   3rd Qu.:20.20   3rd Qu.:16.95
## Max.   :24.000   Max.    :711.0   Max.    :22.00   Max.    :37.97
##          medv
## Min.   : 5.00
## 1st Qu.:17.02
## Median :21.20
## Mean   :22.53
## 3rd Qu.:25.00
## Max.   :50.00
```

Transformando log

```
b$crim = log(b$crim)
```

Fazendo Regressão com todas as variaveis

```
reg.mlt=lm(data=b, medv ~ crim + zn + indus + chas + nox + rm + age + dis + rad + tax + ptratio + lstat,  
summary(reg.mlt)
```

```
##  
## Call:  
## lm(formula = medv ~ crim + zn + indus + chas + nox + rm + age +  
##     dis + rad + tax + ptratio + lstat, data = b)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -14.5196  -2.7591  -0.6185   1.8580  26.8435   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)   42.03833    5.128612   8.197 2.15e-15 ***  
## crim          0.241790    0.279864   0.864 0.388033      
## zn            0.045271    0.014355   3.154 0.001710 **    
## indus         0.016041    0.063182   0.254 0.799692      
## chasbounds river 3.031613    0.879877   3.445 0.000619 ***  
## nox          -18.950940    4.038672  -4.692 3.50e-06 ***  
## rm            3.676786    0.425626   8.639 < 2e-16 ***  
## age           0.002473    0.013590   0.182 0.855696      
## dis          -1.400774    0.202931  -6.903 1.58e-11 ***  
## rad           0.184917    0.077317   2.392 0.017145 *     
## tax          -0.012508    0.003850  -3.249 0.001238 **    
## ptratio      -0.912405    0.134699  -6.774 3.59e-11 ***  
## lstat        -0.591869    0.051245 -11.550 < 2e-16 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 4.86 on 493 degrees of freedom  
## Multiple R-squared:  0.7274, Adjusted R-squared:  0.7208   
## F-statistic: 109.6 on 12 and 493 DF,  p-value: < 2.2e-16
```