Exercicio2.R

adans

2025-04-28

# Questão 1  
  
set.seed(25032004) # COLOQUE O SEU DIA MÊS E ANO DE NASCIMENTO NA SET.SEED  
base = read.csv2("base.csv")  
base1 = base[sample(nrow(base), 800),]  
base1$Local = as.factor(base1$Local)  
# Questão 2  
  
# Modelo de Regressão Múltipla  
modelo = lm(Valor ~ Area + Idade + Energia + Local, data=base1)  
summary(modelo)

##   
## Call:  
## lm(formula = Valor ~ Area + Idade + Energia + Local, data = base1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -29574.2 -6507.2 65.6 6946.3 28467.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7596.149 2424.254 3.133 0.00179 \*\*   
## Area 973.508 16.992 57.293 < 2e-16 \*\*\*  
## Idade -459.662 33.754 -13.618 < 2e-16 \*\*\*  
## Energia 33.212 6.698 4.958 8.70e-07 \*\*\*  
## LocalB -1986.654 963.471 -2.062 0.03954 \*   
## LocalC -3790.853 928.331 -4.084 4.89e-05 \*\*\*  
## LocalD -5655.580 957.707 -5.905 5.22e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9510 on 793 degrees of freedom  
## Multiple R-squared: 0.8204, Adjusted R-squared: 0.819   
## F-statistic: 603.6 on 6 and 793 DF, p-value: < 2.2e-16

# Questão 3  
  
# Analise Grafica da relacao das variaveis independentes com a variavel dependente  
library(car)

## Carregando pacotes exigidos: carData

anova(modelo)

## Analysis of Variance Table  
##   
## Response: Valor  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Area 1 3.0390e+11 3.0390e+11 3360.490 < 2.2e-16 \*\*\*  
## Idade 1 1.8154e+10 1.8154e+10 200.748 < 2.2e-16 \*\*\*  
## Energia 1 1.9803e+09 1.9803e+09 21.898 3.381e-06 \*\*\*  
## Local 3 3.4865e+09 1.1622e+09 12.851 3.321e-08 \*\*\*  
## Residuals 793 7.1713e+10 9.0433e+07   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Questão 4  
  
summary(modelo)

##   
## Call:  
## lm(formula = Valor ~ Area + Idade + Energia + Local, data = base1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -29574.2 -6507.2 65.6 6946.3 28467.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7596.149 2424.254 3.133 0.00179 \*\*   
## Area 973.508 16.992 57.293 < 2e-16 \*\*\*  
## Idade -459.662 33.754 -13.618 < 2e-16 \*\*\*  
## Energia 33.212 6.698 4.958 8.70e-07 \*\*\*  
## LocalB -1986.654 963.471 -2.062 0.03954 \*   
## LocalC -3790.853 928.331 -4.084 4.89e-05 \*\*\*  
## LocalD -5655.580 957.707 -5.905 5.22e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9510 on 793 degrees of freedom  
## Multiple R-squared: 0.8204, Adjusted R-squared: 0.819   
## F-statistic: 603.6 on 6 and 793 DF, p-value: < 2.2e-16

# Questão 5  
  
# Coeficiente de Determinação Ajustado  
modelo = lm(Valor ~ Area + Idade + Energia + Local, data=base1)  
summary(modelo)

##   
## Call:  
## lm(formula = Valor ~ Area + Idade + Energia + Local, data = base1)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -29574.2 -6507.2 65.6 6946.3 28467.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 7596.149 2424.254 3.133 0.00179 \*\*   
## Area 973.508 16.992 57.293 < 2e-16 \*\*\*  
## Idade -459.662 33.754 -13.618 < 2e-16 \*\*\*  
## Energia 33.212 6.698 4.958 8.70e-07 \*\*\*  
## LocalB -1986.654 963.471 -2.062 0.03954 \*   
## LocalC -3790.853 928.331 -4.084 4.89e-05 \*\*\*  
## LocalD -5655.580 957.707 -5.905 5.22e-09 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9510 on 793 degrees of freedom  
## Multiple R-squared: 0.8204, Adjusted R-squared: 0.819   
## F-statistic: 603.6 on 6 and 793 DF, p-value: < 2.2e-16

# Questão 6  
  
# Importancia de cada variavel no modelo  
library(relaimpo)

## Carregando pacotes exigidos: MASS

## Carregando pacotes exigidos: boot

##   
## Anexando pacote: 'boot'

## O seguinte objeto é mascarado por 'package:car':  
##   
## logit

## Carregando pacotes exigidos: survey

## Carregando pacotes exigidos: grid

## Carregando pacotes exigidos: Matrix

## Carregando pacotes exigidos: survival

##   
## Anexando pacote: 'survival'

## O seguinte objeto é mascarado por 'package:boot':  
##   
## aml

##   
## Anexando pacote: 'survey'

## O seguinte objeto é mascarado por 'package:graphics':  
##   
## dotchart

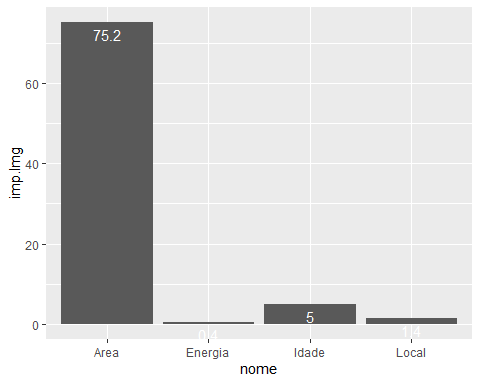
## Carregando pacotes exigidos: mitools

## This is the global version of package relaimpo.

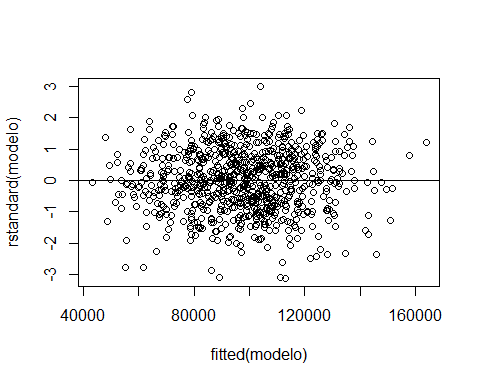
## If you are a non-US user, a version with the interesting additional metric pmvd is available

## from Ulrike Groempings web site at prof.beuth-hochschule.de/groemping.

imp = calc.relimp(modelo)  
var.exp = data.frame(round(imp$lmg\*100,1))  
colnames(var.exp) = "imp.lmg"  
nome = rownames(var.exp)  
var.exp = data.frame(nome,var.exp)  
library(ggplot2)  
ggplot(var.exp,aes(nome,imp.lmg)) +  
 geom\_bar(stat = "identity")+  
 geom\_text(aes(label = imp.lmg), vjust = 1.5, colour = "white")



# Questão 7  
plot(fitted(modelo), rstandard(modelo))  
abline(0,0)



par(mfrow = c(2,2))  
plot(modelo)

