

# dados novos

*Fabio Michel*

*22 de agosto de 2016*

O presente trabalho visa implementar um modelo para a previsão de preços de imóveis para a cidade de Pelotas- RS.

Foram feitas mudanças na base original, onde foi inserido manualmente informações sobre os apartamentos. Foi analisado se os apartamentos eram coberturas ou não. Assim, foi criada uma variável dummie para cobertura.

Inicialmente, Iremos carregar alguns pacotes necessários e também a base de dados.

## carregar pacotes

```
if (!require("pacman")) install.packages("pacman") pacman::p_load(data.table, extremevalues)
```

## Carregar a base de dados

```
data <- read.csv("imoveis2.csv", sep = ";", dec = ",") data <- data.table(data)
```

## Remove some NAs

```
data <- data[area != "NA"] data <- data[,preco := preco/1000] data[,cidade := "Pelotas"] data <- data[,cidade := droplevels(cidade)]
```

## Remove outliers

Assumi que na variável dummie criada não tinha nenhum outliers, pois a base de dados foi avaliada manualmente. E por essa razão não foi inserido a variável cobertura nessa função.

```
outs <- apply(data[,.(preco,quartos,suites,vagas,area)], 2, getOutliers) outs_ind <- list() for (i in 1:length(outs)) { outs_ind[[i]] <- outs[[i]]$iRight } outs_ind <- unique(unlist(outs_ind)) data <- data[!outs_ind] data[,bairro := as.factor(toupper(bairro))]
```

## Classificamos os Apartamentos dos principais bairros na variavel aptos

```
... aptos <- subset(data, tipo == "Apartamento") n_bairro <- aptos[, length(preco), by = bairro]
princ_bairros <- as.character(n_bairro[V1 > 25, bairro]) aptos <- subset(data, tipo == "Apartamento" &
bairro %in% princ_bairros) aptosbairro <- aptos[, droplevels(bairro)] aptostipo <- aptos[, droplevels(tipo)]
aptos$endereco <- aptos[, droplevels(endereco)] ...
```