

Relatório 1 - Regressão

Flavio Margarito Martins de Barros

14/05/2022

Conjunto de dados

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
## Carregando os pacotes
require(readxl)
require(corrplot)
require(psych)
require(kableExtra)
require(caret)
require(GGally)

## Lendo o banco de dados
dados <- read_excel(path = "Concrete_Data.xls", sheet = 1)

## Trocando os nomes das variáveis para o português
colnames(dados) <- c("cimento", "escoria", "cinza", "agua", "super_plastificante",
                     "agregador_grosso", "agregador_fino", "idade", "forca_compressiva")

## Sumario dos dados
```

Preparação dos dados

```
## Separando o conjunto de dados em treino e teste
set.seed(2)
inTrain <- createDataPartition(dados$forca_compressiva, p = 7/10)[[1]]
treino <- dados[inTrain,]
teste <- dados[-inTrain,]

## Mantendo casos completos em treino e teste
treino <- treino[complete.cases(treino),]
teste <- teste[complete.cases(teste),]

## Separando a variavel resposta, categóricas e numericas
resposta <- treino$forca_compressiva
resposta_teste <- teste$forca_compressiva

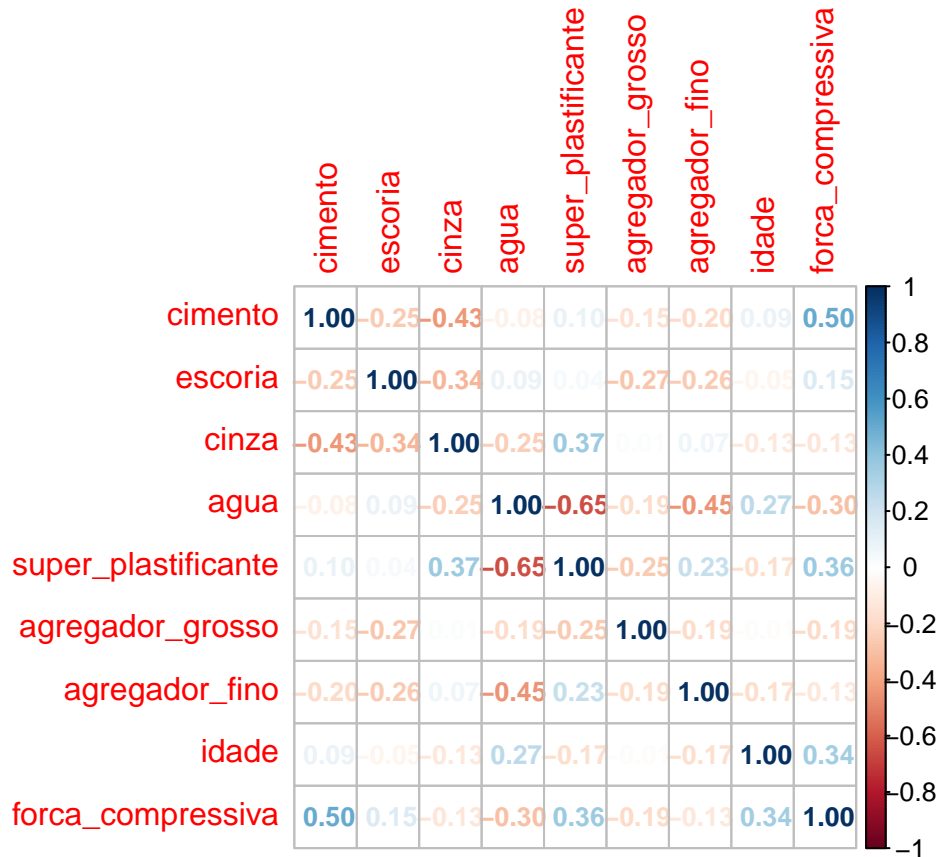
## Retendo as numéricas
Ind_numericas <- colnames(treino)[sapply(treino, is.numeric)]
```

```
Ind_categoricas <- colnames(treino)[sapply(treino, function(x) !is.numeric(x))]
numericas <- treino[,Ind_numericas]
categorias <- treino[,Ind_categoricas]
```

Redução de dimensionalidade

```
## Analisando as correlações
```

```
M <- cor(numericas, use = 'complete.obs')
corrplot(M, method='number', diag = T, number.cex = 0.8)
```



```
summary(M[upper.tri(M)])
```

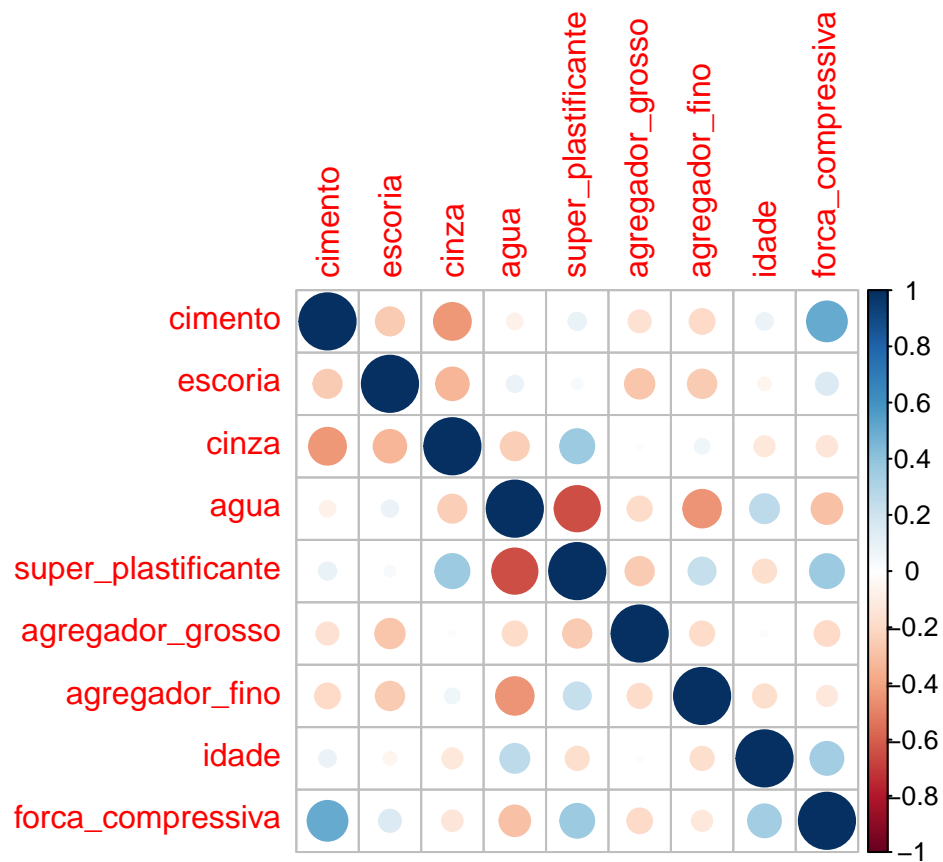
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.64810 -0.24911 -0.13175 -0.07276  0.08627  0.50191
```

```
## Imprimindo as correlações na forma de círculos
```

```
M <- cor(numericas, use = 'complete.obs')
summary(M[upper.tri(M)])
```

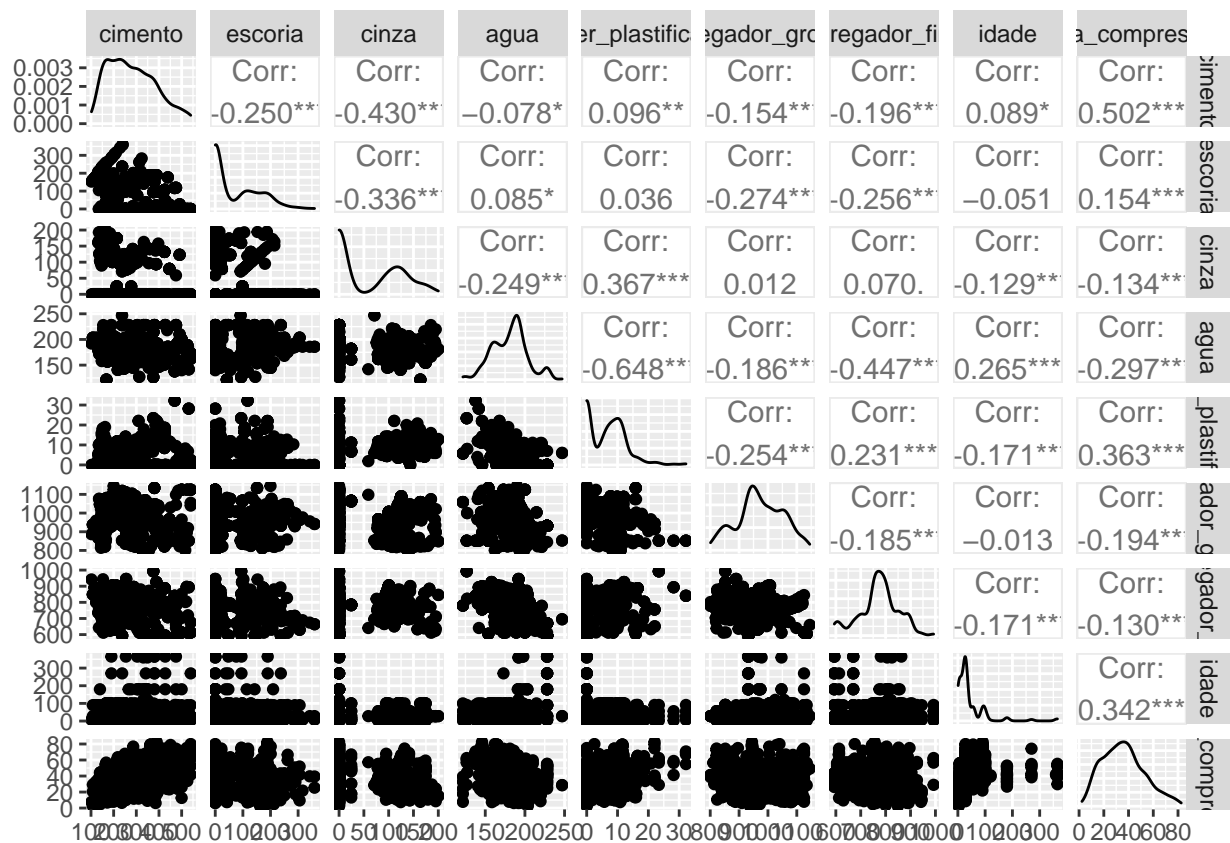
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## -0.64810 -0.24911 -0.13175 -0.07276  0.08627  0.50191
```

```
corrplot(M, method='circle')
```



```
save(M, file = 'M.rda')

## Visualizando as correlações
ggpairs(numericas)
```



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.