Aprendizagem Estatística em Altas Dimensões [MAE0501/MAE5904/IBI5904]

Ícaro Maia Santos de Castro[1]
Rayssa de Carvalho Roberto[2]
Rodrigo Aoyama Nakahara[3]
Rodrigo Araujo[4]
Vitor Hugo Vieira de Lima[5]

Novembro de 2020

Sumário

[1]Número USP: 11866921
[2]Número USP: 10940828
[3]Número USP: 3510922
[4]Número USP: 9299208
[5]Número USP: 10263886

Importando os dados / Limpando / Inspecionando

```
diabetes <- read_csv("C:\\Users\\Rodrigo Araujo\\Documents\\IME-USP\\Aprendizagem Estatística em Altas D
```

```
## Parsed with column specification:
## cols(
##   Pregnancies = col_double(),
##   Glucose = col_double(),
##   BloodPressure = col_double(),
##   SkinThickness = col_double(),
##   Insulin = col_double(),
##   BMI = col_double(),
##   DiabetesPedigreeFunction = col_double(),
##   Age = col_double(),
##   Outcome = col_double()
## )
```

```
#diabetes <- read.csv("diabetes.csv")
head(diabetes) %>% kable(caption="Dados.")
```

Tabela 1: Dados.

Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
6	148	72	35	0	33,6	0,63	50	1
1	85	66	29	0	26,6	0,35	31	0
8	183	64	0	0	23,3	0,67	32	1
1	89	66	23	94	28,1	0,17	21	0
0	137	40	35	168	43,1	2,29	33	1
5	116	74	0	0	25,6	0,20	30	0

```
dim(diabetes)
```

```
## [1] 768   9
```

Renomear coluna e ajustar níveis de categórica

```
diabetes[, 2:6][diabetes[, 2:6] == 0] <- NA

colnames(diabetes)[9] <- "diabetes"

diabetes$diabetes <- as.factor(diabetes$diabetes)

levels(diabetes$diabetes) <- c("No","Yes")
```

Visualização dos Dados

Estrutura dos Dados

```
str(diabetes)
```

```
## tibble [768 x 9] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Pregnancies             : num [1:768] 6 1 8 1 0 5 3 10 2 8 ...
##  $ Glucose                 : num [1:768] 148 85 183 89 137 116 78 115 197 125 ...
##  $ BloodPressure           : num [1:768] 72 66 64 66 40 74 50 NA 70 96 ...
##  $ SkinThickness           : num [1:768] 35 29 NA 23 35 NA 32 NA 45 NA ...
##  $ Insulin                 : num [1:768] NA NA NA 94 168 NA 88 NA 543 NA ...
##  $ BMI                     : num [1:768] 33.6 26.6 23.3 28.1 43.1 25.6 31 35.3 30.5 NA ...
##  $ DiabetesPedigreeFunction: num [1:768] 0.627 0.351 0.672 0.167 2.288 ...
##  $ Age                     : num [1:768] 50 31 32 21 33 30 26 29 53 54 ...
##  $ diabetes                : Factor w/ 2 levels "No","Yes": 2 1 2 1 2 1 2 1 2 2 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   Pregnancies = col_double(),
##   ..   Glucose = col_double(),
##   ..   BloodPressure = col_double(),
##   ..   SkinThickness = col_double(),
##   ..   Insulin = col_double(),
##   ..   BMI = col_double(),
##   ..   DiabetesPedigreeFunction = col_double(),
##   ..   Age = col_double(),
##   ..   Outcome = col_double()
##   .. )
```

train / test

```
# para reprodução
set.seed(23)
nrows <- nrow(diabetes)
index <- sample(1:nrows, 0.7 * nrows)   # shuffle and divide
# train <- diab                         # 768 test data (100%)
train <- diabetes[index,]                   # 537 test data (70%)
test <- diabetes[-index,]                   # 231 test data (30%)
```

Proporção de diabetes (Benign / Malignant)

train

```
prop.table(table(train$diabetes))
```

```
##
##        No       Yes
## 0.6405959 0.3594041
```

test

```
prop.table(table(test$diabetes))
```

```
##
##        No       Yes
## 0.6753247 0.3246753
```

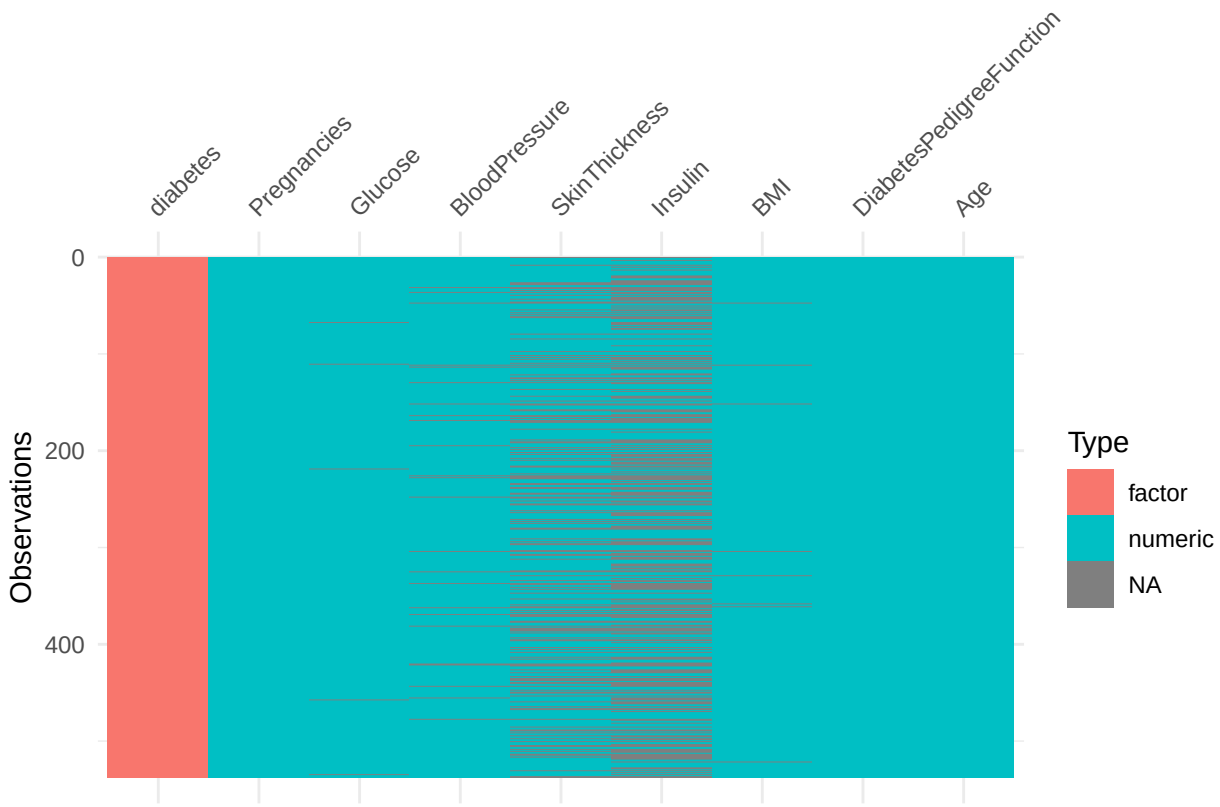
Analise dos Missings

Descritiva inicial

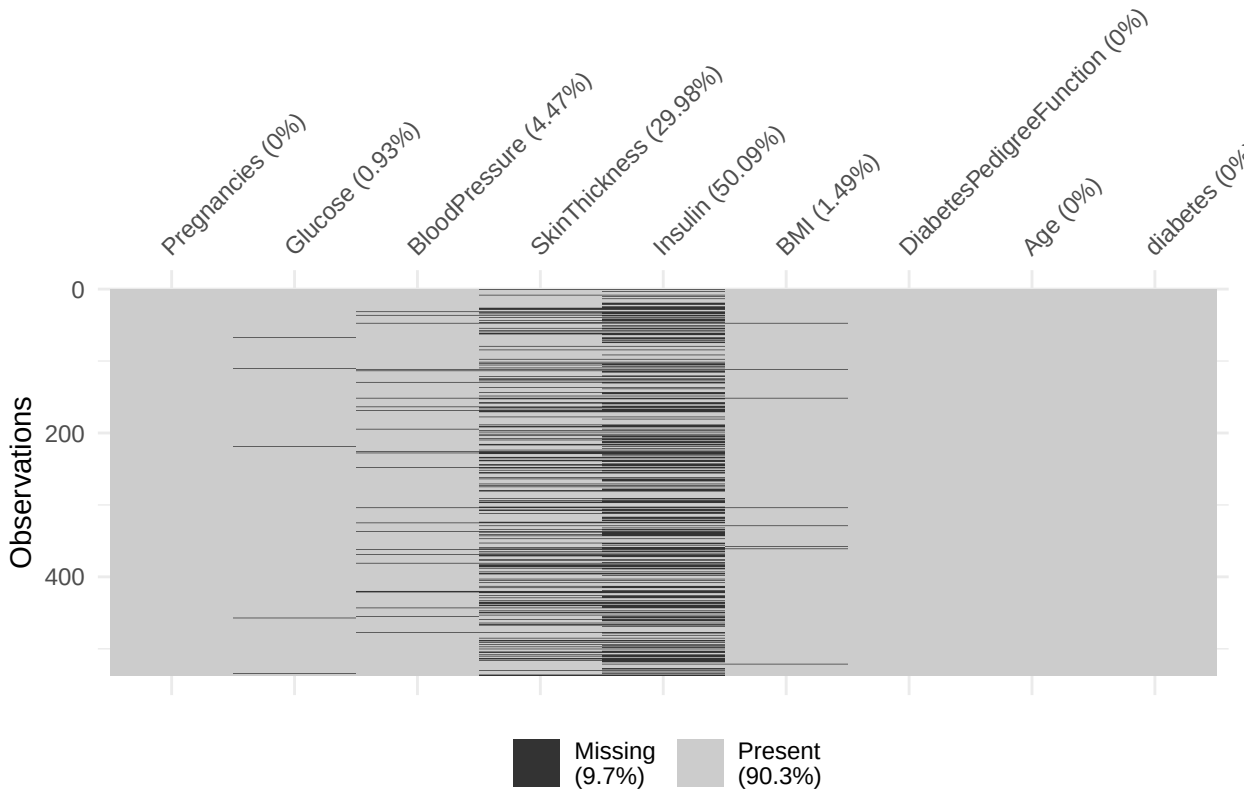
```
summary(train)
```

```
##   Pregnancies        Glucose      BloodPressure    SkinThickness
##  Min.   : 0.000   Min.   : 44.0   Min.   : 24.00   Min.   : 7.00
##  1st Qu.: 1.000   1st Qu.:100.0   1st Qu.: 64.00   1st Qu.:22.75
##  Median : 3.000   Median :117.0   Median : 72.00   Median :30.00
##  Mean   : 3.903   Mean   :122.7   Mean   : 72.72   Mean   :29.68
##  3rd Qu.: 6.000   3rd Qu.:143.2   3rd Qu.: 80.00   3rd Qu.:37.00
##  Max.   :15.000   Max.   :199.0   Max.   :114.00   Max.   :63.00
##                   NA's   :5       NA's   :24       NA's   :161
##     Insulin           BMI        DiabetesPedigreeFunction      Age
##  Min.   : 14.0   Min.   :18.20   Min.   :0.078            Min.   :21.00
##  1st Qu.: 76.0   1st Qu.:27.60   1st Qu.:0.238            1st Qu.:24.00
##  Median :126.0   Median :32.40   Median :0.361            Median :29.00
##  Mean   :157.1   Mean   :32.55   Mean   :0.463            Mean   :33.56
##  3rd Qu.:190.2   3rd Qu.:36.80   3rd Qu.:0.619            3rd Qu.:41.00
##  Max.   :846.0   Max.   :67.10   Max.   :2.420            Max.   :72.00
##  NA's   :269     NA's   :8
##  diabetes
##  No :344
##  Yes:193
##
##
##
##
##
```

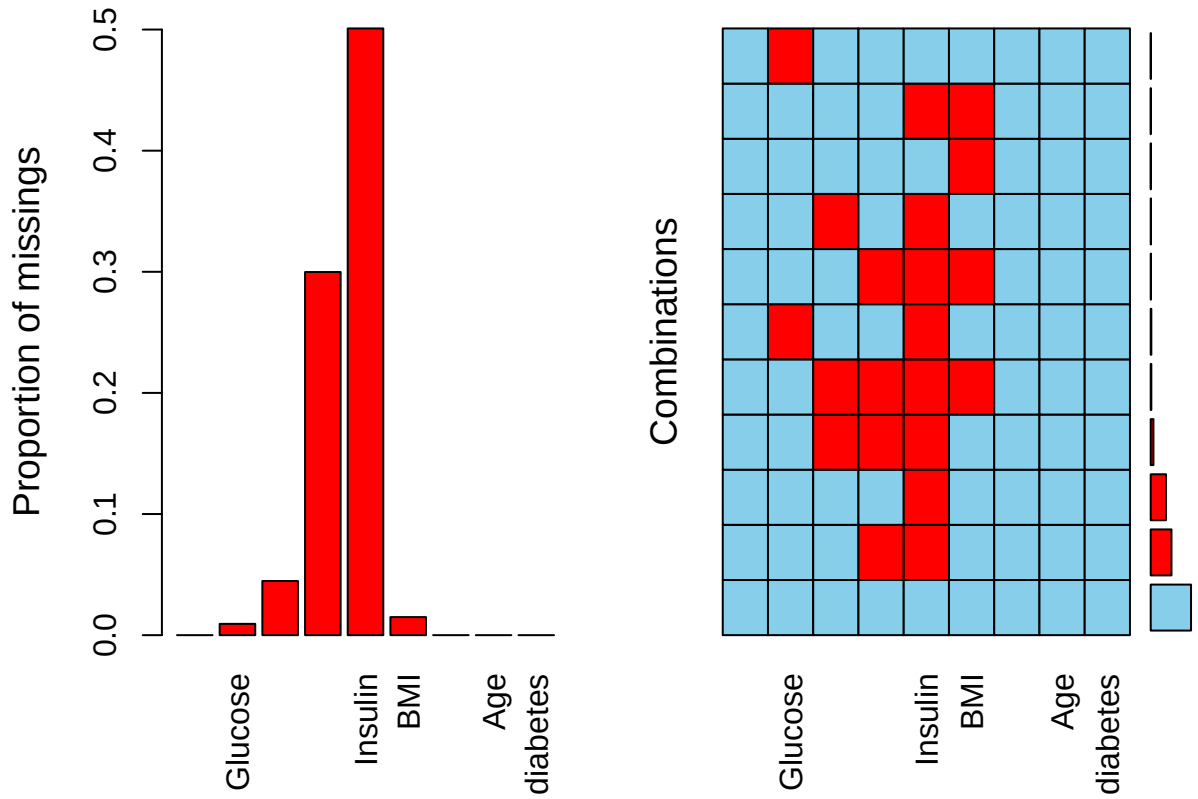
```
vis_dat(train)
```



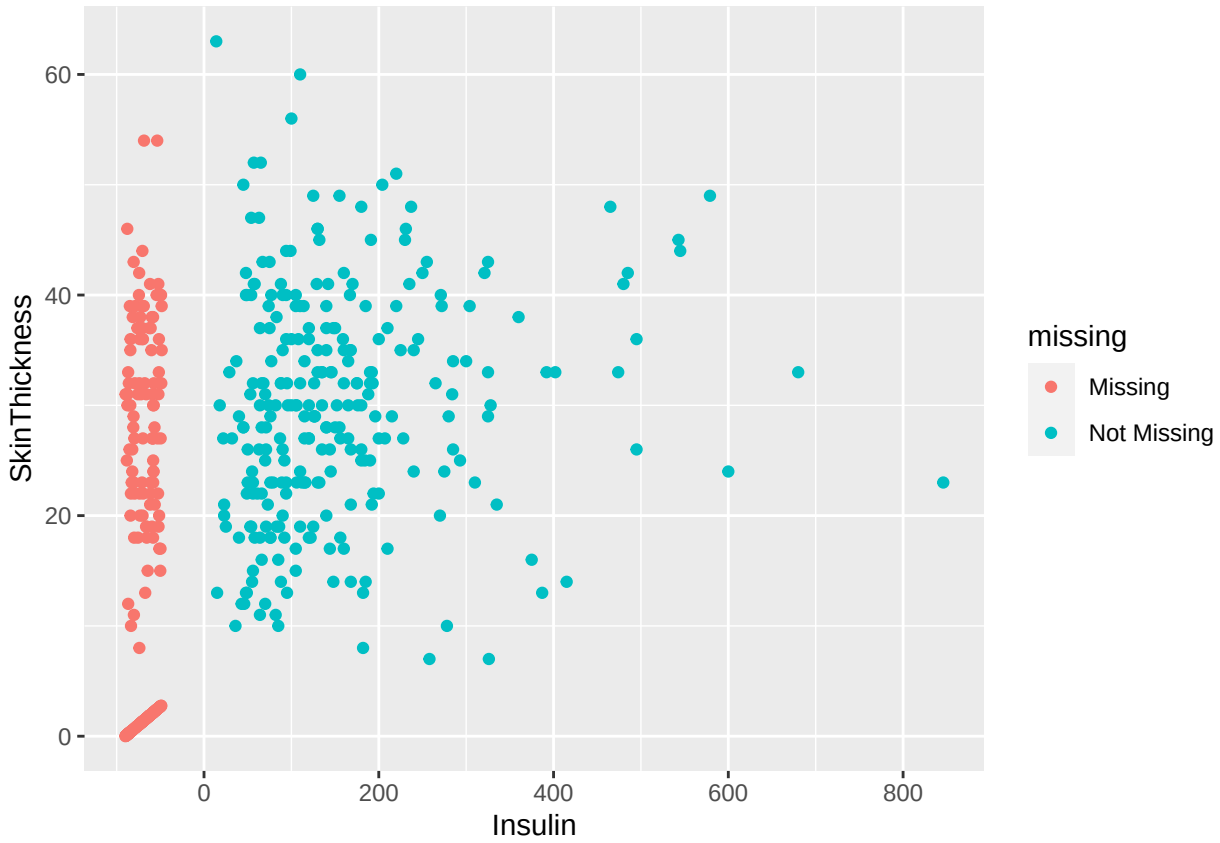

```
vis_miss(train)
```



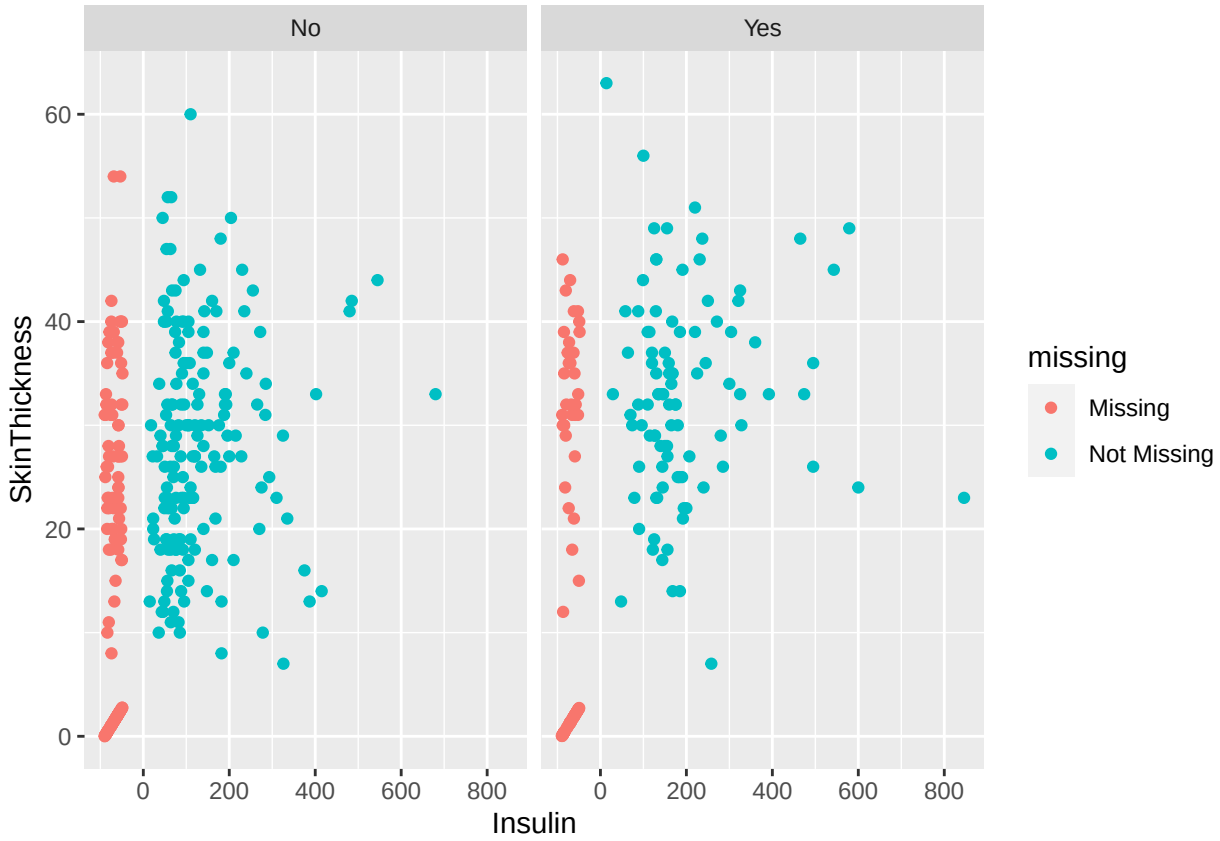

```
aggr(train) # Missings têm padrões?
```



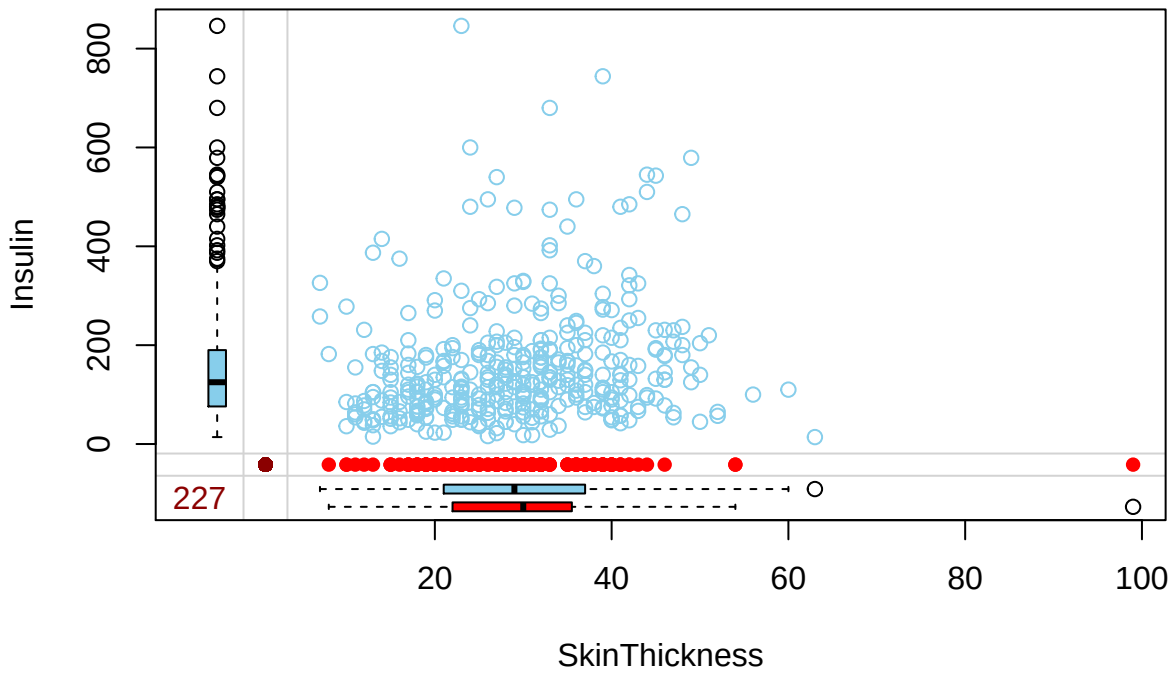

```
ggplot(train, aes(x = Insulin, y = SkinThickness)) + # Padrão de missing entre 2 vars com mais missings
  geom_miss_point()
```




```
ggplot(train, aes(x = Insulin, y = SkinThickness)) + # Padrão de missing entre 2 vars com mais missings
  geom_miss_point() + facet_wrap(~ diabetes)
```




```
marginplot(diabetes[c(4,5)])
```



Possibilidades de imputação

```
library(mice)
```

```
## Warning: package 'mice' was built under R version 4.0.3
```

```
##
## Attaching package: 'mice'
```

```
## The following objects are masked from 'package:base':
##
##     cbind, rbind
```

```
imputacao_train <- mice(data = train , m = 5, maxit = 50, meth = 'pmm', seed = 25)
```

```
##
##  iter imp variable
##   1   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   1   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   1   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   1   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   1   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   2   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   2   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   2   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   2   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   2   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   3   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   3   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   3   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   3   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   3   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   4   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   4   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   4   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   4   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   4   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   5   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   5   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   5   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   5   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   5   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   6   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   6   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   6   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   6   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   6   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   7   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   7   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   7   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   7   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   7   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   8   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   8   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   8   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   8   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   8   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   9   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   9   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   9   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   9   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   9   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  10   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  10   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  10   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  10   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  10   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  11   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  11   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  11   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  11   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  11   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  12   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##  12   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   12   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   12   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   12   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   13   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   13   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   13   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   13   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   13   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   14   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   14   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   14   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   14   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   14   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   15   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   15   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   15   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   15   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   15   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   16   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   16   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   16   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   16   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   16   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   17   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   17   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   17   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   17   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   17   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   18   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   18   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   18   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   18   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   18   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   19   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   19   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   19   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   19   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   19   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   20   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   20   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   20   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   20   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   20   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   21   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   21   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   21   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   21   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   21   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   22   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   22   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   22   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   22   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   22   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   23   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   23   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   23   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   23   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   23   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   24   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   24   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   24   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   24   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   24   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   25   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   25   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   25   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   25   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   25   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   26   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   26   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   26   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   26   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   26   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   27   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   27   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   27   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   27   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   27   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   28   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   28   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   28   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   28   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   28   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   29   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   29   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   29   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   29   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   29   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   30   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   30   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   30   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   30   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   30   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   31   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   31   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   31   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   31   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   31   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   32   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   32   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   32   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   32   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   32   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   33   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   33   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   33   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   33   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   33   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   34   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   34   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   34   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   34   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   34   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   35   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   35   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   35   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   35   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   35   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   36   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   36   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   36   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   36   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   36   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   37   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   37   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   37   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   37   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   37   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   38   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   38   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   38   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   38   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   38   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   39   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   39   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   39   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   39   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   39   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   40   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   40   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   40   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   40   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   40   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   41   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   41   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   41   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   41   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   41   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   42   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   42   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   42   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   42   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   42   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   43   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   43   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   43   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   43   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   43   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   44   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   44   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   44   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   44   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

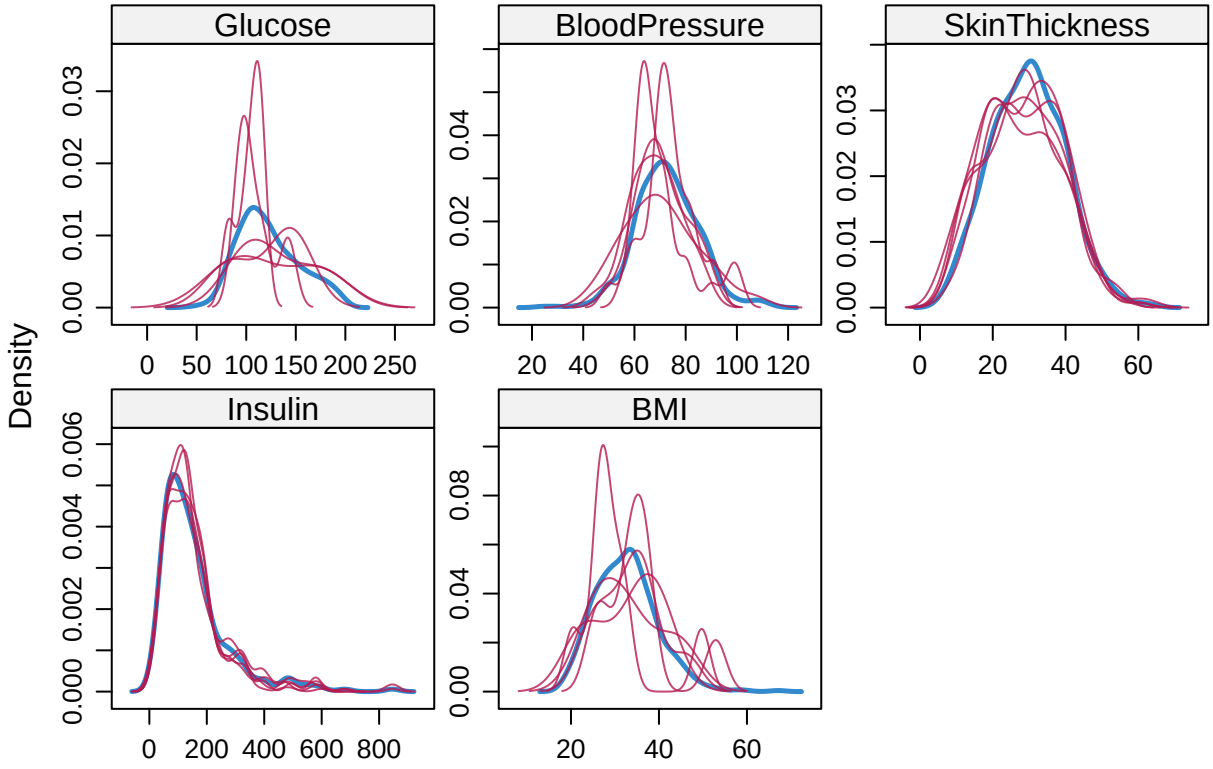
```
##   44   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   45   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   45   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   45   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   45   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   45   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   46   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   46   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   46   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   46   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   46   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   47   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   47   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   47   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   47   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   47   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   48   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   48   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   48   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   48   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   48   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   49   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   49   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   49   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   49   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   49   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   50   1  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   50   2  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   50   3  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   50   4  Glucose  BloodPressure  SkinThickness  Insulin  BMI
##   50   5  Glucose  BloodPressure  SkinThickness  Insulin  BMI
```

```
summary(imputacao_train)
```

```
## Class: mids
## Number of multiple imputations:  5
## Imputation methods:
##              Pregnancies                  Glucose            BloodPressure
##                       ""                    "pmm"                    "pmm"
##            SkinThickness                  Insulin                      BMI
##                    "pmm"                    "pmm"                    "pmm"
## DiabetesPedigreeFunction                      Age                 diabetes
##                       ""                     ""                       ""
## PredictorMatrix:
##               Pregnancies Glucose BloodPressure SkinThickness Insulin BMI
## Pregnancies             0       1             1             1       1   1
## Glucose                 1       0             1             1       1   1
## BloodPressure           1       1             0             1       1   1
## SkinThickness           1       1             1             0       1   1
## Insulin                 1       1             1             1       0   1
## BMI                     1       1             1             1       1   0
##               DiabetesPedigreeFunction Age diabetes
## Pregnancies                          1   1        1
```

```
## Glucose                              1   1        1
## BloodPressure                        1   1        1
## SkinThickness                        1   1        1
## Insulin                              1   1        1
## BMI                                  1   1        1
```

```
imputado_train_1 <- complete(imputacao_train, 1)
imputado_train_2 <- complete(imputacao_train, 2)
imputado_train_3 <- complete(imputacao_train, 3)
imputado_train_4 <- complete(imputacao_train, 4)
imputado_train_5 <- complete(imputacao_train, 5)

library(lattice)
densityplot(imputacao_train)
```




```
train$Glucose <- apply(cbind(imputado_train_1$Glucose, imputado_train_2$Glucose, imputado_train_3$Glucos

train$BloodPressure <- apply(cbind(imputado_train_1$BloodPressure, imputado_train_2$BloodPressure, imput

train$SkinThickness <- apply(cbind(imputado_train_1$SkinThickness, imputado_train_2$SkinThickness, imput

train$Insulin <- apply(cbind(imputado_train_1$Insulin, imputado_train_2$Insulin, imputado_train_3$Insuli

train$BMI <- apply(cbind(imputado_train_1$BMI, imputado_train_2$BMI, imputado_train_3$BMI, imputado_trai
```

```
imputacao_test <- mice(data = test , m = 5, maxit = 50, meth = 'pmm', seed = 25)
```

```
##
##  iter imp variable
##   1   1  BloodPressure  SkinThickness  Insulin  BMI
##   1   2  BloodPressure  SkinThickness  Insulin  BMI
##   1   3  BloodPressure  SkinThickness  Insulin  BMI
##   1   4  BloodPressure  SkinThickness  Insulin  BMI
##   1   5  BloodPressure  SkinThickness  Insulin  BMI
##   2   1  BloodPressure  SkinThickness  Insulin  BMI
##   2   2  BloodPressure  SkinThickness  Insulin  BMI
##   2   3  BloodPressure  SkinThickness  Insulin  BMI
##   2   4  BloodPressure  SkinThickness  Insulin  BMI
##   2   5  BloodPressure  SkinThickness  Insulin  BMI
##   3   1  BloodPressure  SkinThickness  Insulin  BMI
##   3   2  BloodPressure  SkinThickness  Insulin  BMI
##   3   3  BloodPressure  SkinThickness  Insulin  BMI
##   3   4  BloodPressure  SkinThickness  Insulin  BMI
##   3   5  BloodPressure  SkinThickness  Insulin  BMI
##   4   1  BloodPressure  SkinThickness  Insulin  BMI
##   4   2  BloodPressure  SkinThickness  Insulin  BMI
##   4   3  BloodPressure  SkinThickness  Insulin  BMI
##   4   4  BloodPressure  SkinThickness  Insulin  BMI
##   4   5  BloodPressure  SkinThickness  Insulin  BMI
##   5   1  BloodPressure  SkinThickness  Insulin  BMI
##   5   2  BloodPressure  SkinThickness  Insulin  BMI
##   5   3  BloodPressure  SkinThickness  Insulin  BMI
##   5   4  BloodPressure  SkinThickness  Insulin  BMI
##   5   5  BloodPressure  SkinThickness  Insulin  BMI
##   6   1  BloodPressure  SkinThickness  Insulin  BMI
##   6   2  BloodPressure  SkinThickness  Insulin  BMI
##   6   3  BloodPressure  SkinThickness  Insulin  BMI
##   6   4  BloodPressure  SkinThickness  Insulin  BMI
##   6   5  BloodPressure  SkinThickness  Insulin  BMI
##   7   1  BloodPressure  SkinThickness  Insulin  BMI
##   7   2  BloodPressure  SkinThickness  Insulin  BMI
##   7   3  BloodPressure  SkinThickness  Insulin  BMI
##   7   4  BloodPressure  SkinThickness  Insulin  BMI
##   7   5  BloodPressure  SkinThickness  Insulin  BMI
##   8   1  BloodPressure  SkinThickness  Insulin  BMI
##   8   2  BloodPressure  SkinThickness  Insulin  BMI
##   8   3  BloodPressure  SkinThickness  Insulin  BMI
##   8   4  BloodPressure  SkinThickness  Insulin  BMI
##   8   5  BloodPressure  SkinThickness  Insulin  BMI
##   9   1  BloodPressure  SkinThickness  Insulin  BMI
##   9   2  BloodPressure  SkinThickness  Insulin  BMI
##   9   3  BloodPressure  SkinThickness  Insulin  BMI
##   9   4  BloodPressure  SkinThickness  Insulin  BMI
##   9   5  BloodPressure  SkinThickness  Insulin  BMI
##   10   1  BloodPressure  SkinThickness  Insulin  BMI
##   10   2  BloodPressure  SkinThickness  Insulin  BMI
##   10   3  BloodPressure  SkinThickness  Insulin  BMI
##   10   4  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   10   5  BloodPressure  SkinThickness  Insulin  BMI
##   11   1  BloodPressure  SkinThickness  Insulin  BMI
##   11   2  BloodPressure  SkinThickness  Insulin  BMI
##   11   3  BloodPressure  SkinThickness  Insulin  BMI
##   11   4  BloodPressure  SkinThickness  Insulin  BMI
##   11   5  BloodPressure  SkinThickness  Insulin  BMI
##   12   1  BloodPressure  SkinThickness  Insulin  BMI
##   12   2  BloodPressure  SkinThickness  Insulin  BMI
##   12   3  BloodPressure  SkinThickness  Insulin  BMI
##   12   4  BloodPressure  SkinThickness  Insulin  BMI
##   12   5  BloodPressure  SkinThickness  Insulin  BMI
##   13   1  BloodPressure  SkinThickness  Insulin  BMI
##   13   2  BloodPressure  SkinThickness  Insulin  BMI
##   13   3  BloodPressure  SkinThickness  Insulin  BMI
##   13   4  BloodPressure  SkinThickness  Insulin  BMI
##   13   5  BloodPressure  SkinThickness  Insulin  BMI
##   14   1  BloodPressure  SkinThickness  Insulin  BMI
##   14   2  BloodPressure  SkinThickness  Insulin  BMI
##   14   3  BloodPressure  SkinThickness  Insulin  BMI
##   14   4  BloodPressure  SkinThickness  Insulin  BMI
##   14   5  BloodPressure  SkinThickness  Insulin  BMI
##   15   1  BloodPressure  SkinThickness  Insulin  BMI
##   15   2  BloodPressure  SkinThickness  Insulin  BMI
##   15   3  BloodPressure  SkinThickness  Insulin  BMI
##   15   4  BloodPressure  SkinThickness  Insulin  BMI
##   15   5  BloodPressure  SkinThickness  Insulin  BMI
##   16   1  BloodPressure  SkinThickness  Insulin  BMI
##   16   2  BloodPressure  SkinThickness  Insulin  BMI
##   16   3  BloodPressure  SkinThickness  Insulin  BMI
##   16   4  BloodPressure  SkinThickness  Insulin  BMI
##   16   5  BloodPressure  SkinThickness  Insulin  BMI
##   17   1  BloodPressure  SkinThickness  Insulin  BMI
##   17   2  BloodPressure  SkinThickness  Insulin  BMI
##   17   3  BloodPressure  SkinThickness  Insulin  BMI
##   17   4  BloodPressure  SkinThickness  Insulin  BMI
##   17   5  BloodPressure  SkinThickness  Insulin  BMI
##   18   1  BloodPressure  SkinThickness  Insulin  BMI
##   18   2  BloodPressure  SkinThickness  Insulin  BMI
##   18   3  BloodPressure  SkinThickness  Insulin  BMI
##   18   4  BloodPressure  SkinThickness  Insulin  BMI
##   18   5  BloodPressure  SkinThickness  Insulin  BMI
##   19   1  BloodPressure  SkinThickness  Insulin  BMI
##   19   2  BloodPressure  SkinThickness  Insulin  BMI
##   19   3  BloodPressure  SkinThickness  Insulin  BMI
##   19   4  BloodPressure  SkinThickness  Insulin  BMI
##   19   5  BloodPressure  SkinThickness  Insulin  BMI
##   20   1  BloodPressure  SkinThickness  Insulin  BMI
##   20   2  BloodPressure  SkinThickness  Insulin  BMI
##   20   3  BloodPressure  SkinThickness  Insulin  BMI
##   20   4  BloodPressure  SkinThickness  Insulin  BMI
##   20   5  BloodPressure  SkinThickness  Insulin  BMI
##   21   1  BloodPressure  SkinThickness  Insulin  BMI
##   21   2  BloodPressure  SkinThickness  Insulin  BMI
##   21   3  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   21   4  BloodPressure  SkinThickness  Insulin  BMI
##   21   5  BloodPressure  SkinThickness  Insulin  BMI
##   22   1  BloodPressure  SkinThickness  Insulin  BMI
##   22   2  BloodPressure  SkinThickness  Insulin  BMI
##   22   3  BloodPressure  SkinThickness  Insulin  BMI
##   22   4  BloodPressure  SkinThickness  Insulin  BMI
##   22   5  BloodPressure  SkinThickness  Insulin  BMI
##   23   1  BloodPressure  SkinThickness  Insulin  BMI
##   23   2  BloodPressure  SkinThickness  Insulin  BMI
##   23   3  BloodPressure  SkinThickness  Insulin  BMI
##   23   4  BloodPressure  SkinThickness  Insulin  BMI
##   23   5  BloodPressure  SkinThickness  Insulin  BMI
##   24   1  BloodPressure  SkinThickness  Insulin  BMI
##   24   2  BloodPressure  SkinThickness  Insulin  BMI
##   24   3  BloodPressure  SkinThickness  Insulin  BMI
##   24   4  BloodPressure  SkinThickness  Insulin  BMI
##   24   5  BloodPressure  SkinThickness  Insulin  BMI
##   25   1  BloodPressure  SkinThickness  Insulin  BMI
##   25   2  BloodPressure  SkinThickness  Insulin  BMI
##   25   3  BloodPressure  SkinThickness  Insulin  BMI
##   25   4  BloodPressure  SkinThickness  Insulin  BMI
##   25   5  BloodPressure  SkinThickness  Insulin  BMI
##   26   1  BloodPressure  SkinThickness  Insulin  BMI
##   26   2  BloodPressure  SkinThickness  Insulin  BMI
##   26   3  BloodPressure  SkinThickness  Insulin  BMI
##   26   4  BloodPressure  SkinThickness  Insulin  BMI
##   26   5  BloodPressure  SkinThickness  Insulin  BMI
##   27   1  BloodPressure  SkinThickness  Insulin  BMI
##   27   2  BloodPressure  SkinThickness  Insulin  BMI
##   27   3  BloodPressure  SkinThickness  Insulin  BMI
##   27   4  BloodPressure  SkinThickness  Insulin  BMI
##   27   5  BloodPressure  SkinThickness  Insulin  BMI
##   28   1  BloodPressure  SkinThickness  Insulin  BMI
##   28   2  BloodPressure  SkinThickness  Insulin  BMI
##   28   3  BloodPressure  SkinThickness  Insulin  BMI
##   28   4  BloodPressure  SkinThickness  Insulin  BMI
##   28   5  BloodPressure  SkinThickness  Insulin  BMI
##   29   1  BloodPressure  SkinThickness  Insulin  BMI
##   29   2  BloodPressure  SkinThickness  Insulin  BMI
##   29   3  BloodPressure  SkinThickness  Insulin  BMI
##   29   4  BloodPressure  SkinThickness  Insulin  BMI
##   29   5  BloodPressure  SkinThickness  Insulin  BMI
##   30   1  BloodPressure  SkinThickness  Insulin  BMI
##   30   2  BloodPressure  SkinThickness  Insulin  BMI
##   30   3  BloodPressure  SkinThickness  Insulin  BMI
##   30   4  BloodPressure  SkinThickness  Insulin  BMI
##   30   5  BloodPressure  SkinThickness  Insulin  BMI
##   31   1  BloodPressure  SkinThickness  Insulin  BMI
##   31   2  BloodPressure  SkinThickness  Insulin  BMI
##   31   3  BloodPressure  SkinThickness  Insulin  BMI
##   31   4  BloodPressure  SkinThickness  Insulin  BMI
##   31   5  BloodPressure  SkinThickness  Insulin  BMI
##   32   1  BloodPressure  SkinThickness  Insulin  BMI
##   32   2  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   32   3  BloodPressure  SkinThickness  Insulin  BMI
##   32   4  BloodPressure  SkinThickness  Insulin  BMI
##   32   5  BloodPressure  SkinThickness  Insulin  BMI
##   33   1  BloodPressure  SkinThickness  Insulin  BMI
##   33   2  BloodPressure  SkinThickness  Insulin  BMI
##   33   3  BloodPressure  SkinThickness  Insulin  BMI
##   33   4  BloodPressure  SkinThickness  Insulin  BMI
##   33   5  BloodPressure  SkinThickness  Insulin  BMI
##   34   1  BloodPressure  SkinThickness  Insulin  BMI
##   34   2  BloodPressure  SkinThickness  Insulin  BMI
##   34   3  BloodPressure  SkinThickness  Insulin  BMI
##   34   4  BloodPressure  SkinThickness  Insulin  BMI
##   34   5  BloodPressure  SkinThickness  Insulin  BMI
##   35   1  BloodPressure  SkinThickness  Insulin  BMI
##   35   2  BloodPressure  SkinThickness  Insulin  BMI
##   35   3  BloodPressure  SkinThickness  Insulin  BMI
##   35   4  BloodPressure  SkinThickness  Insulin  BMI
##   35   5  BloodPressure  SkinThickness  Insulin  BMI
##   36   1  BloodPressure  SkinThickness  Insulin  BMI
##   36   2  BloodPressure  SkinThickness  Insulin  BMI
##   36   3  BloodPressure  SkinThickness  Insulin  BMI
##   36   4  BloodPressure  SkinThickness  Insulin  BMI
##   36   5  BloodPressure  SkinThickness  Insulin  BMI
##   37   1  BloodPressure  SkinThickness  Insulin  BMI
##   37   2  BloodPressure  SkinThickness  Insulin  BMI
##   37   3  BloodPressure  SkinThickness  Insulin  BMI
##   37   4  BloodPressure  SkinThickness  Insulin  BMI
##   37   5  BloodPressure  SkinThickness  Insulin  BMI
##   38   1  BloodPressure  SkinThickness  Insulin  BMI
##   38   2  BloodPressure  SkinThickness  Insulin  BMI
##   38   3  BloodPressure  SkinThickness  Insulin  BMI
##   38   4  BloodPressure  SkinThickness  Insulin  BMI
##   38   5  BloodPressure  SkinThickness  Insulin  BMI
##   39   1  BloodPressure  SkinThickness  Insulin  BMI
##   39   2  BloodPressure  SkinThickness  Insulin  BMI
##   39   3  BloodPressure  SkinThickness  Insulin  BMI
##   39   4  BloodPressure  SkinThickness  Insulin  BMI
##   39   5  BloodPressure  SkinThickness  Insulin  BMI
##   40   1  BloodPressure  SkinThickness  Insulin  BMI
##   40   2  BloodPressure  SkinThickness  Insulin  BMI
##   40   3  BloodPressure  SkinThickness  Insulin  BMI
##   40   4  BloodPressure  SkinThickness  Insulin  BMI
##   40   5  BloodPressure  SkinThickness  Insulin  BMI
##   41   1  BloodPressure  SkinThickness  Insulin  BMI
##   41   2  BloodPressure  SkinThickness  Insulin  BMI
##   41   3  BloodPressure  SkinThickness  Insulin  BMI
##   41   4  BloodPressure  SkinThickness  Insulin  BMI
##   41   5  BloodPressure  SkinThickness  Insulin  BMI
##   42   1  BloodPressure  SkinThickness  Insulin  BMI
##   42   2  BloodPressure  SkinThickness  Insulin  BMI
##   42   3  BloodPressure  SkinThickness  Insulin  BMI
##   42   4  BloodPressure  SkinThickness  Insulin  BMI
##   42   5  BloodPressure  SkinThickness  Insulin  BMI
##   43   1  BloodPressure  SkinThickness  Insulin  BMI
```

```
##   43   2  BloodPressure  SkinThickness  Insulin  BMI
##   43   3  BloodPressure  SkinThickness  Insulin  BMI
##   43   4  BloodPressure  SkinThickness  Insulin  BMI
##   43   5  BloodPressure  SkinThickness  Insulin  BMI
##   44   1  BloodPressure  SkinThickness  Insulin  BMI
##   44   2  BloodPressure  SkinThickness  Insulin  BMI
##   44   3  BloodPressure  SkinThickness  Insulin  BMI
##   44   4  BloodPressure  SkinThickness  Insulin  BMI
##   44   5  BloodPressure  SkinThickness  Insulin  BMI
##   45   1  BloodPressure  SkinThickness  Insulin  BMI
##   45   2  BloodPressure  SkinThickness  Insulin  BMI
##   45   3  BloodPressure  SkinThickness  Insulin  BMI
##   45   4  BloodPressure  SkinThickness  Insulin  BMI
##   45   5  BloodPressure  SkinThickness  Insulin  BMI
##   46   1  BloodPressure  SkinThickness  Insulin  BMI
##   46   2  BloodPressure  SkinThickness  Insulin  BMI
##   46   3  BloodPressure  SkinThickness  Insulin  BMI
##   46   4  BloodPressure  SkinThickness  Insulin  BMI
##   46   5  BloodPressure  SkinThickness  Insulin  BMI
##   47   1  BloodPressure  SkinThickness  Insulin  BMI
##   47   2  BloodPressure  SkinThickness  Insulin  BMI
##   47   3  BloodPressure  SkinThickness  Insulin  BMI
##   47   4  BloodPressure  SkinThickness  Insulin  BMI
##   47   5  BloodPressure  SkinThickness  Insulin  BMI
##   48   1  BloodPressure  SkinThickness  Insulin  BMI
##   48   2  BloodPressure  SkinThickness  Insulin  BMI
##   48   3  BloodPressure  SkinThickness  Insulin  BMI
##   48   4  BloodPressure  SkinThickness  Insulin  BMI
##   48   5  BloodPressure  SkinThickness  Insulin  BMI
##   49   1  BloodPressure  SkinThickness  Insulin  BMI
##   49   2  BloodPressure  SkinThickness  Insulin  BMI
##   49   3  BloodPressure  SkinThickness  Insulin  BMI
##   49   4  BloodPressure  SkinThickness  Insulin  BMI
##   49   5  BloodPressure  SkinThickness  Insulin  BMI
##   50   1  BloodPressure  SkinThickness  Insulin  BMI
##   50   2  BloodPressure  SkinThickness  Insulin  BMI
##   50   3  BloodPressure  SkinThickness  Insulin  BMI
##   50   4  BloodPressure  SkinThickness  Insulin  BMI
##   50   5  BloodPressure  SkinThickness  Insulin  BMI
```

```
summary(imputacao_test)
```

```
## Class: mids
## Number of multiple imputations:  5
## Imputation methods:
##              Pregnancies                  Glucose            BloodPressure
##                       ""                       ""                   "pmm"
##            SkinThickness                  Insulin                      BMI
##                    "pmm"                    "pmm"                   "pmm"
## DiabetesPedigreeFunction                      Age                 diabetes
##                       ""                       ""                       ""
## PredictorMatrix:
##               Pregnancies Glucose BloodPressure SkinThickness Insulin BMI
```

```
## Pregnancies             0       1             1             1       1   1
## Glucose                 1       0             1             1       1   1
## BloodPressure           1       1             0             1       1   1
## SkinThickness           1       1             1             0       1   1
## Insulin                 1       1             1             1       0   1
## BMI                     1       1             1             1       1   0
##               DiabetesPedigreeFunction Age diabetes
## Pregnancies                          1   1        1
## Glucose                              1   1        1
## BloodPressure                        1   1        1
## SkinThickness                        1   1        1
## Insulin                              1   1        1
## BMI                                  1   1        1
```

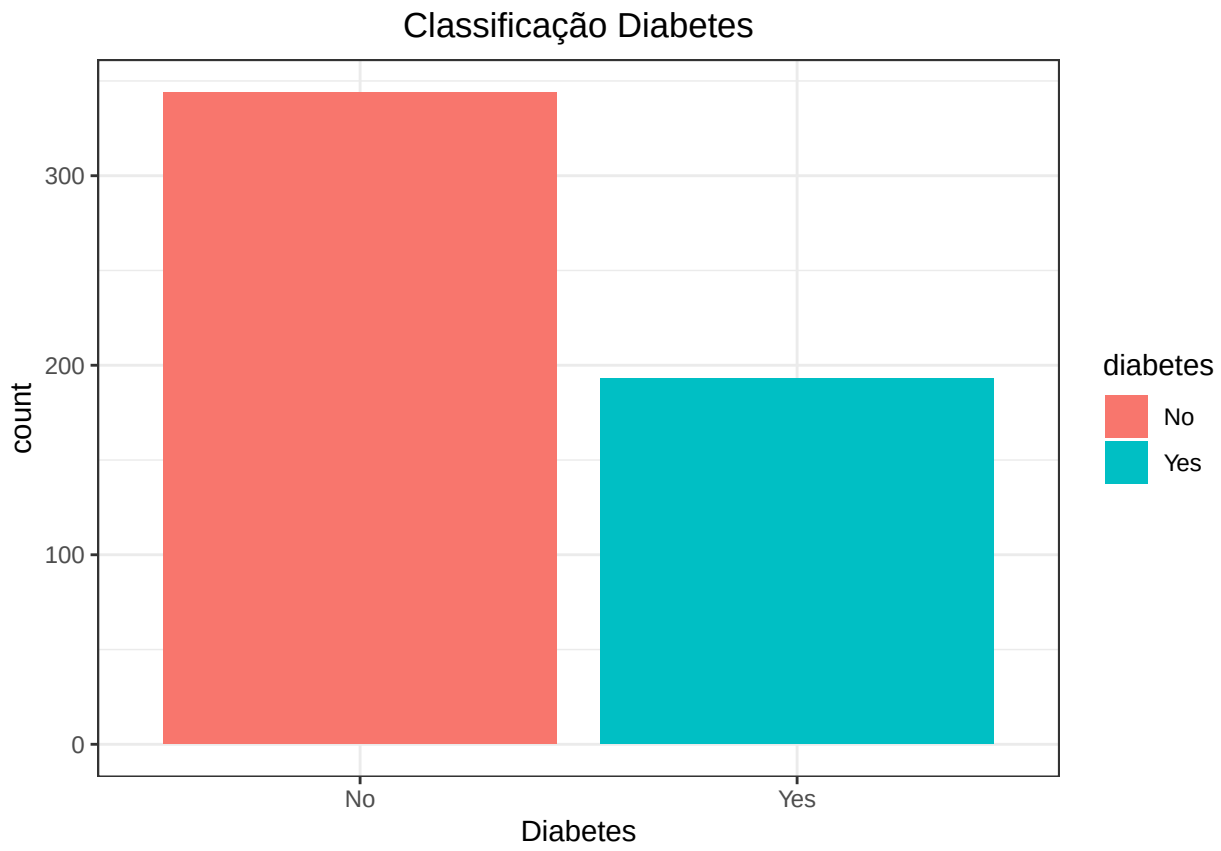
```
imputado_test_1 <- complete(imputacao_test, 1)
imputado_test_2 <- complete(imputacao_test, 2)
imputado_test_3 <- complete(imputacao_test, 3)
imputado_test_4 <- complete(imputacao_test, 4)
imputado_test_5 <- complete(imputacao_test, 5)
```

```
test$Glucose <- apply(cbind(imputado_test_1$Glucose, imputado_test_2$Glucose, imputado_test_3$Glucose, i

test$BloodPressure <- apply(cbind(imputado_test_1$BloodPressure, imputado_test_2$BloodPressure, imputado

test$SkinThickness <- apply(cbind(imputado_test_1$SkinThickness, imputado_test_2$SkinThickness, imputado

test$Insulin <- apply(cbind(imputado_test_1$Insulin, imputado_test_2$Insulin, imputado_test_3$Insulin, i

test$BMI <- apply(cbind(imputado_test_1$BMI, imputado_test_2$BMI, imputado_test_3$BMI, imputado_test_4$I
```

Análise Descritiva

Distribuição da variável Diabetes

```
ggplot(train, aes(diabetes, fill = diabetes)) +
  geom_bar() +
  theme_bw() +
  labs(title = "Classificação Diabetes", x = "Diabetes") +
  theme(plot.title = element_text(hjust = 0.5))
```



Correlação entre cada variável

```
library(PerformanceAnalytics)
```

```
## Loading required package: xts
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

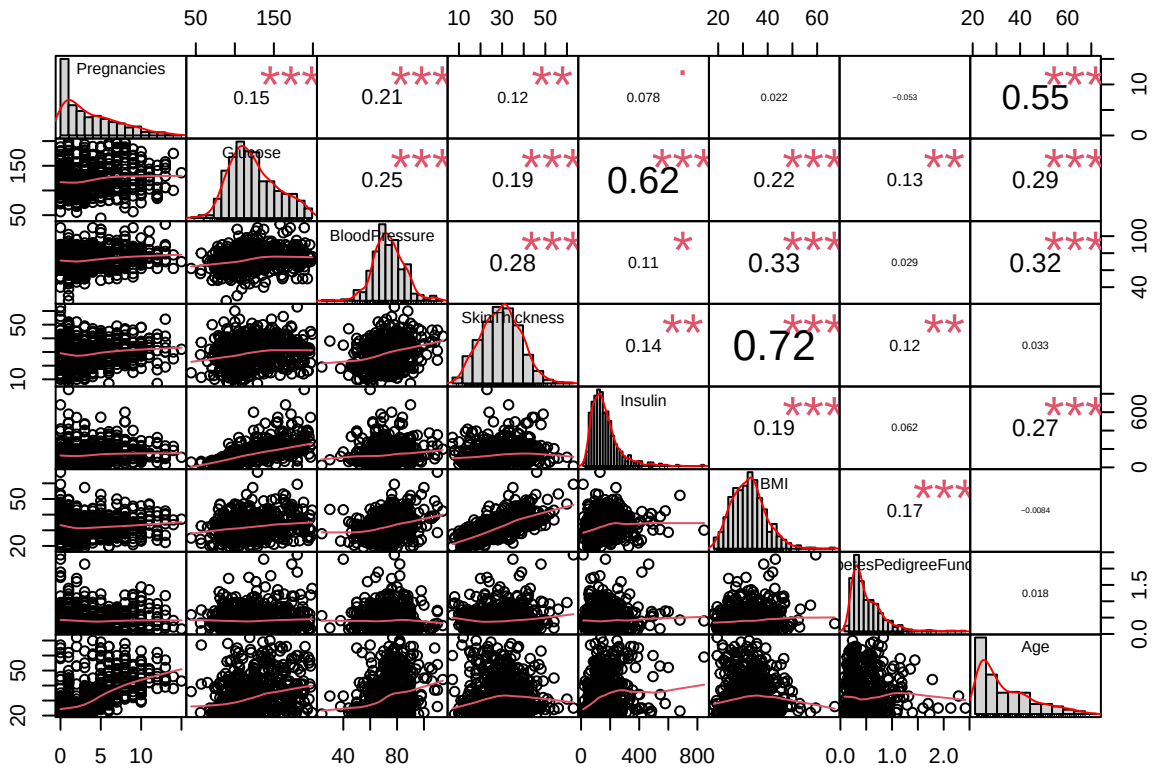
```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
##
## Attaching package: 'xts'
```

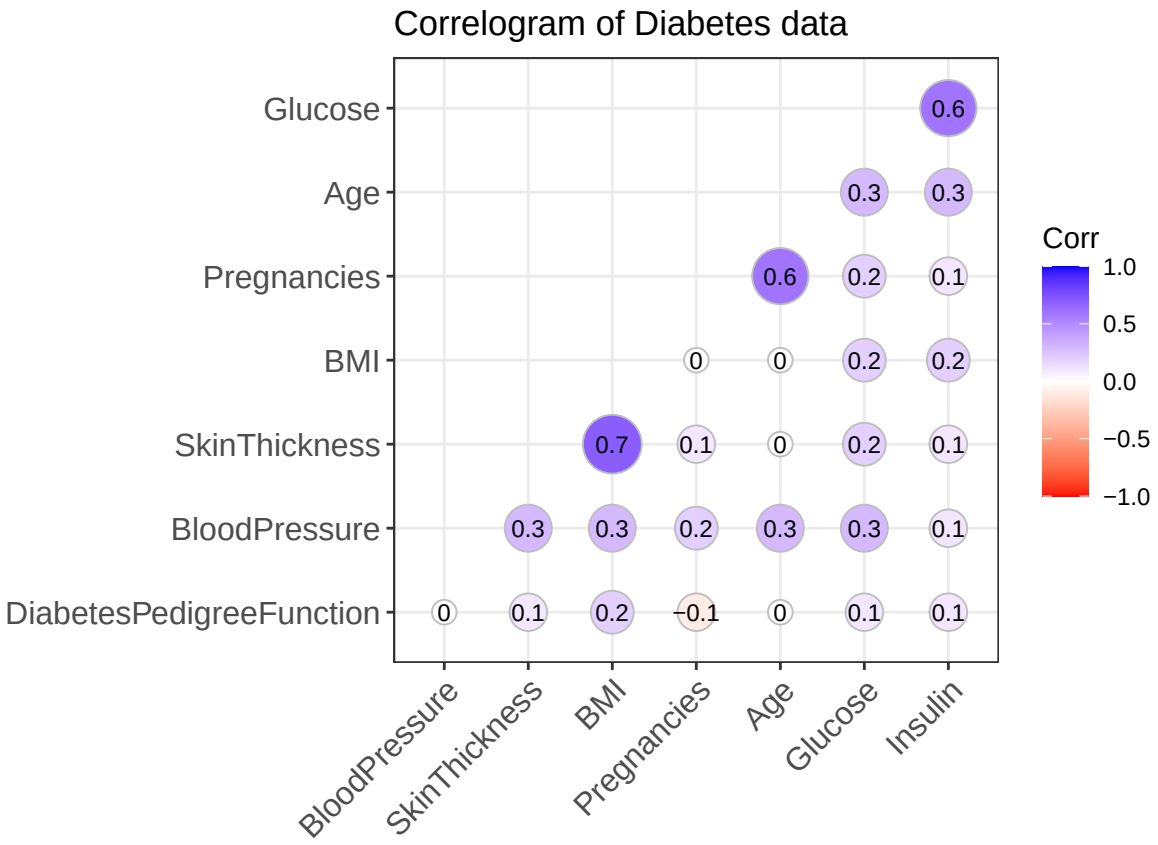
```
## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```
##
## Attaching package: 'PerformanceAnalytics'


## The following object is masked from 'package:graphics':
##
##     legend
```

```
chart.Correlation(train[,-9], histogram=TRUE, col="grey10", pch=1, main="Correlação entre ás variáveis
```




```
library(ggcorrplot)

corr<-round(cor(train[,-9]),1)
ggcorrplot(corr, hc.order = TRUE,
           type = "lower",
           lab = TRUE,
           lab_size = 3,
           method="circle",
           colors = c("red", "white", "blue"),
           title="Correlogram of Diabetes data",
           ggtheme=theme_bw)
```



Modelagem

SVM

```
library(caret)
```

```
##
## Attaching package: 'caret'

## The following object is masked from 'package:vegan':
##
##     tolerance

## The following objects are masked from 'package:MLmetrics':
##
##     MAE, RMSE

## The following object is masked from 'package:survival':
##
##     cluster
```

```
## The following object is masked from 'package:purrr':
##
##     lift
```

```
library(e1071)
```

```
##
## Attaching package: 'e1071'
```

```
## The following objects are masked from 'package:PerformanceAnalytics':
##
##     kurtosis, skewness
```

```
## The following object is masked from 'package:Hmisc':
##
##     impute
```

```
set.seed(123)

linear.tune <- e1071::tune.svm(diabetes ~.,
                               data   = train,
                               kernel = 'linear',
                               cost   = c(0.001, 0.01, 0.1, 1, 5, 10))
summary(linear.tune)
```

```
##
## Parameter tuning of 'svm':
##
## - sampling method: 10-fold cross validation
##
## - best parameters:
##  cost
##  0.01
##
## - best performance: 0.2254018
##
## - Detailed performance results:
##    cost     error dispersion
## 1 1e-03 0.3591894 0.07461689
## 2 1e-02 0.2254018 0.05260323
## 3 1e-01 0.2291405 0.05587773
## 4 1e+00 0.2272886 0.05437335
## 5 5e+00 0.2366177 0.05092700
## 6 1e+01 0.2366177 0.05092700
```

Matriz de Confusão

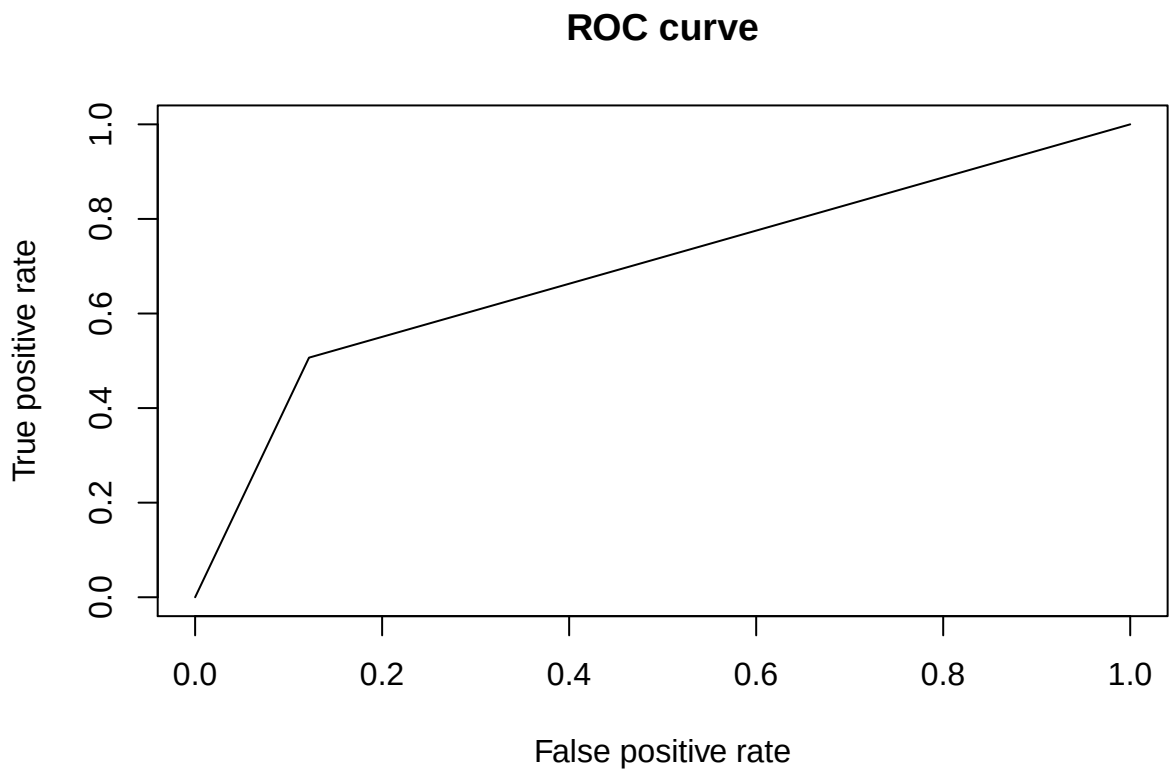
```
svm.real <- test$diabetes

best.linear <- linear.tune$best.model
```

```r
tune.test <- predict(best.linear, test[,-9])

caret::confusionMatrix(data = tune.test,
                       reference = svm.real,
                       positive  = 'Yes')
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  No Yes
##        No  137  37
##        Yes  19  38
##
##                Accuracy : 0.7576
##                  95% CI : (0.697, 0.8114)
##     No Information Rate : 0.6753
##     P-Value [Acc > NIR] : 0.003953
##
##                   Kappa : 0.4104
##
##  Mcnemar's Test P-Value : 0.023103
##
##             Sensitivity : 0.5067
##             Specificity : 0.8782
##          Pos Pred Value : 0.6667
##          Neg Pred Value : 0.7874
##              Prevalence : 0.3247
##          Detection Rate : 0.1645
##    Detection Prevalence : 0.2468
##       Balanced Accuracy : 0.6924
##
##        'Positive' Class : Yes
##
```

Curva ROC

```r
svm.predobj <- ROCR::prediction(predictions = as.numeric(x = tune.test ),
                                labels      = as.numeric(x = svm.real))
svm.perform <- ROCR::performance(prediction.obj = svm.predobj,
                                 measure        = 'tpr',
                                 x.measure      = 'fpr')
plot(x = svm.perform, main = 'ROC curve')
```




```
MLmetrics::F1_Score(y_pred   = tune.test ,
                    y_true   = svm.real,
                    positive = "Yes"); pROC::auc(response = as.numeric(x = svm.real),
                                                 predictor = as.numeric(x = tune.test))
```

```
## [1] 0.5757576
```

```
## Setting levels: control = 1, case = 2
```

```
## Setting direction: controls < cases
```

```
## Area under the curve: 0.6924
```

RandomForest

```
library(randomForest)
```

```
## randomForest 4.6-14
```

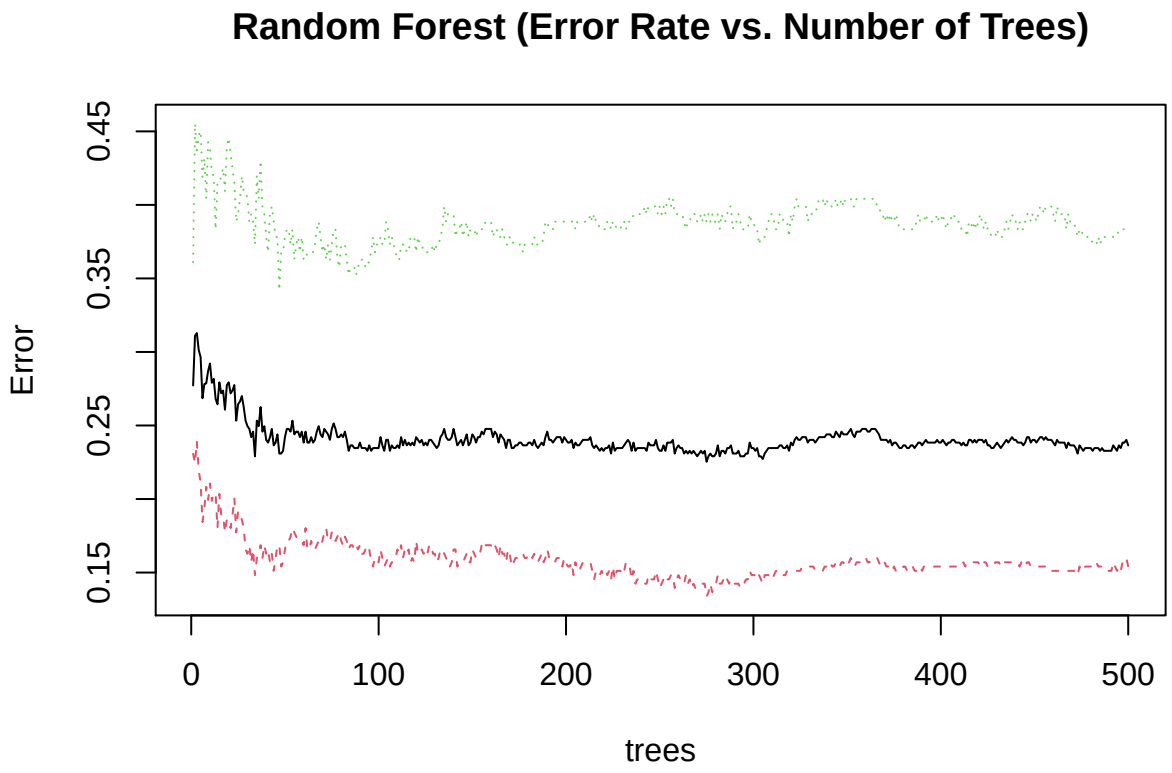
```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'


## The following object is masked from 'package:psych':
##
##     outlier


## The following object is masked from 'package:dplyr':
##
##     combine


## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
learn_rf <- randomForest(diabetes~., data=train, ntree=500, proximity=T, importance=T)

pre_rf   <- predict(learn_rf, test[,-9])

cm_rf    <- confusionMatrix(pre_rf, test$diabetes)

cm_rf
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction  No Yes
##        No  130  32
##        Yes  26  43
##
##                Accuracy : 0.7489
##                  95% CI : (0.6878, 0.8035)
##     No Information Rate : 0.6753
##     P-Value [Acc > NIR] : 0.009121
##
##                   Kappa : 0.4153
##
##  Mcnemar's Test P-Value : 0.511482
##
##             Sensitivity : 0.8333
##             Specificity : 0.5733
##          Pos Pred Value : 0.8025
##          Neg Pred Value : 0.6232
##              Prevalence : 0.6753
##          Detection Rate : 0.5628
##    Detection Prevalence : 0.7013
##       Balanced Accuracy : 0.7033
##
##        'Positive' Class : No
##
```

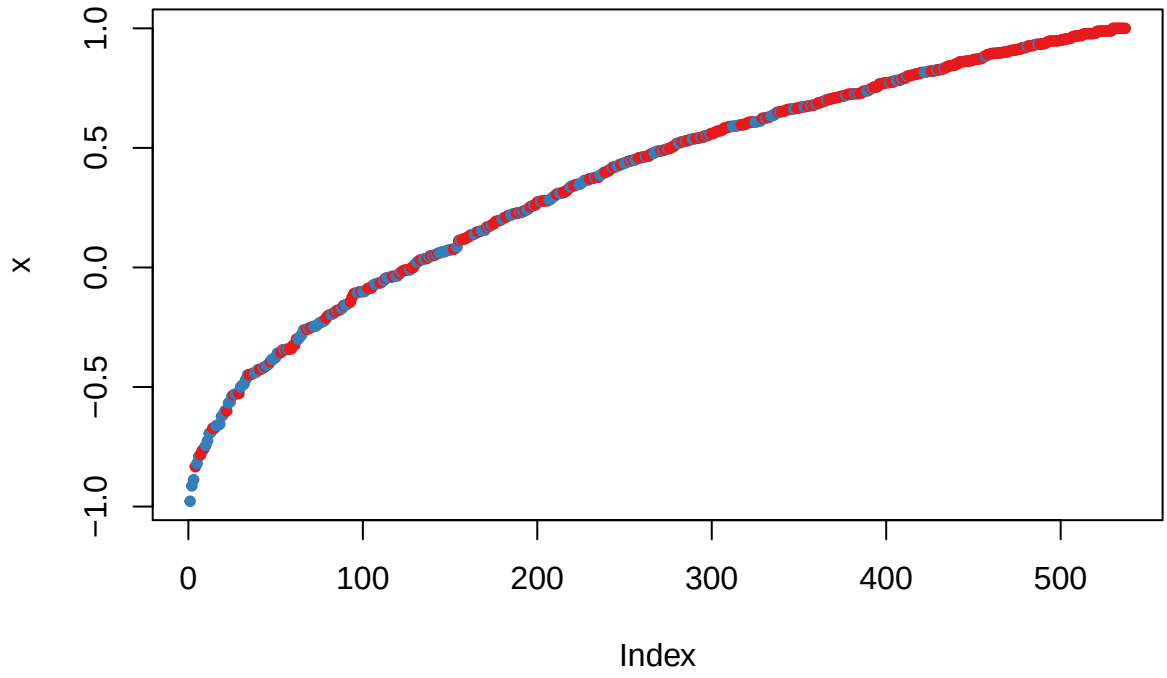
```
plot(learn_rf, main="Random Forest (Error Rate vs. Number of Trees)")
```



Prediction Plot

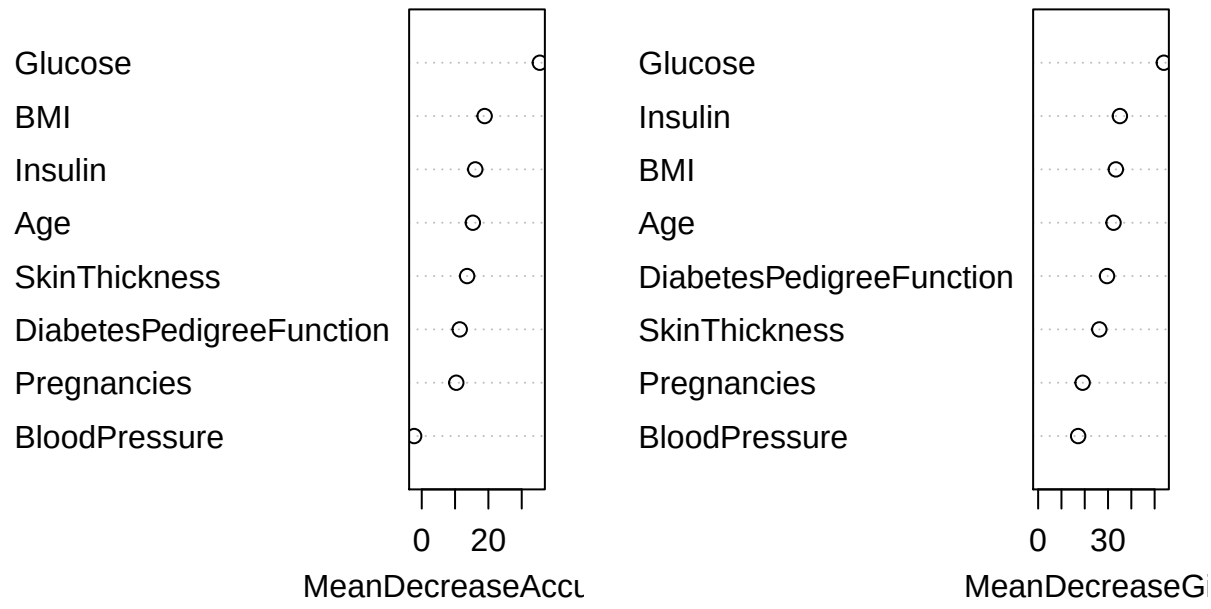
```
plot(margin(learn_rf,test$diabetes))
```

```
## Warning in RColorBrewer::brewer.pal(nlevs, "Set1"): minimal value for n is 3, returning requested pal
```



Variance Importance Plot

```
varImpPlot(learn_rf)
```


learn_rf
Glucose
BMI
Insulin
Age
SkinThickness
DiabetesPedigreeFunction
Pregnancies
BloodPressure
0
20
MeanDecreaseAccu
Glucose
Insulin
BMI
Age
DiabetesPedigreeFunction
SkinThickness
Pregnancies
BloodPressure
0
30
MeanDecreaseGi