DATASET

kaggle

Search

Create

Home

Competitions

Datasets

Models

Code

Discussions

Learn

More

Your Work

RECENTLY VIEWED

View Active Events

CHAU NGUYỄN ĐẶNG HUỲNH · UPDATED 3 MONTHS AGO

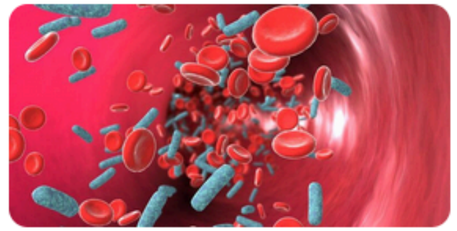
29

New Notebook

Download (13 kB)

Sepsis

This is just a dataset for practicing to classify the sepsis illness only



Data Card | Code (9) | Discussion (0)

About Dataset

Dataset Description:
The provided data is a modified version of a publicly available data source, and is subject to copyright.

Donor of database:

```
The Johns Hopkins University
Johns Hopkins Road
Laurel, MD 20707
(301) 953-6231
```

Usability
8.82

License
CC0: Public Domain

Expected update frequency

VARIÁVEIS

- ID: number to represent patient ID
- PRG: Plasma glucose
- PL: Blood Work Result-1 (mu U/ml)
- PR: Blood Pressure (mm Hg)
- SK: Serum potassium (mm)
- TS: Trasnferrin (mu U/ml)
- M11: Body mass index (weight in kg/(height in m)^2
- BD2: Beta-defensin 2 (mu U/ml)
- Age: patients age (years)
- Insurance: If a patient holds a valid insurance card
- Sepsis: Positive: if a patient in ICU will develop a sepsis , and Negative: otherwise

MÉTODOS

- RStudio
  - Tratamento dos dados
  - Estatística descritiva
  - Teste de normalidade
  - Testes de comparação entre duas medidas
  - Teste de associação

github.com/juliaradula/sepsis_stats

HIPÓTESES

- Existe associação entre desfecho clínico e seguro saúde?
- Entre os pacientes dos grupos Sepse positivo e Sepse negativo quais variáveis é possível ver diferença?
- Quais variáveis associam-se entre si?

TRATAMENTO DOS DADOS

Inicialmente:

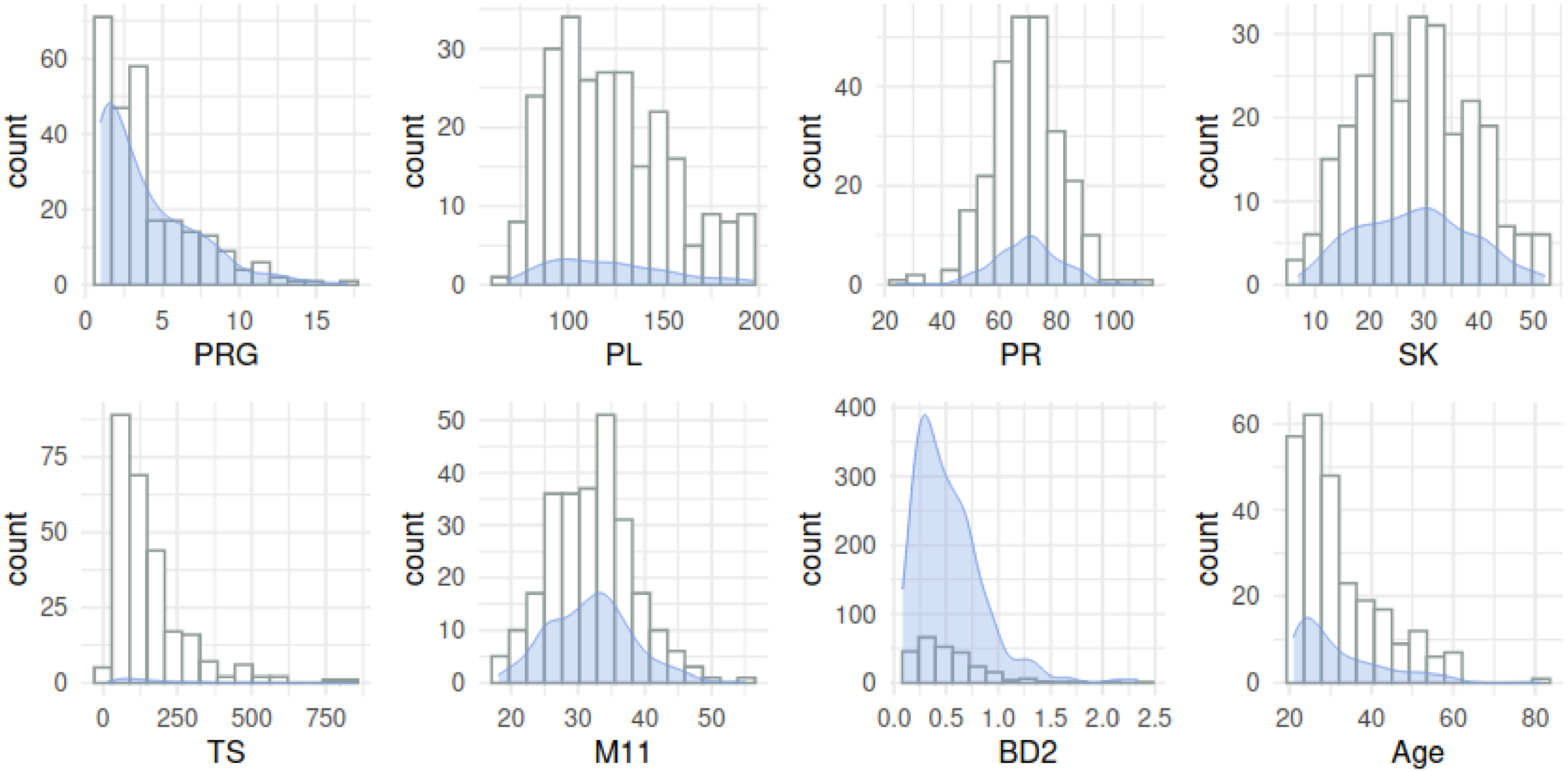
- 599 observações

Após filtrar zeros:

- 261 observações

edit: Isso não se faz!!!!

DISTRIBUIÇÃO DAS VARIÁVEIS NUMÉRICAS



TESTE DE NORMALIDADE

> shapiro.test(dados$variavel)

PRG: W = 0.8, p-value = 2e-15

PL: W = 1, p-value = 5e-07

PR: W = 1, p-value = 0.008

SK: W = 1, p-value = 0.005

TS: W = 0.8, p-value <2e-16

M11: W = 1, p-value = 0.04

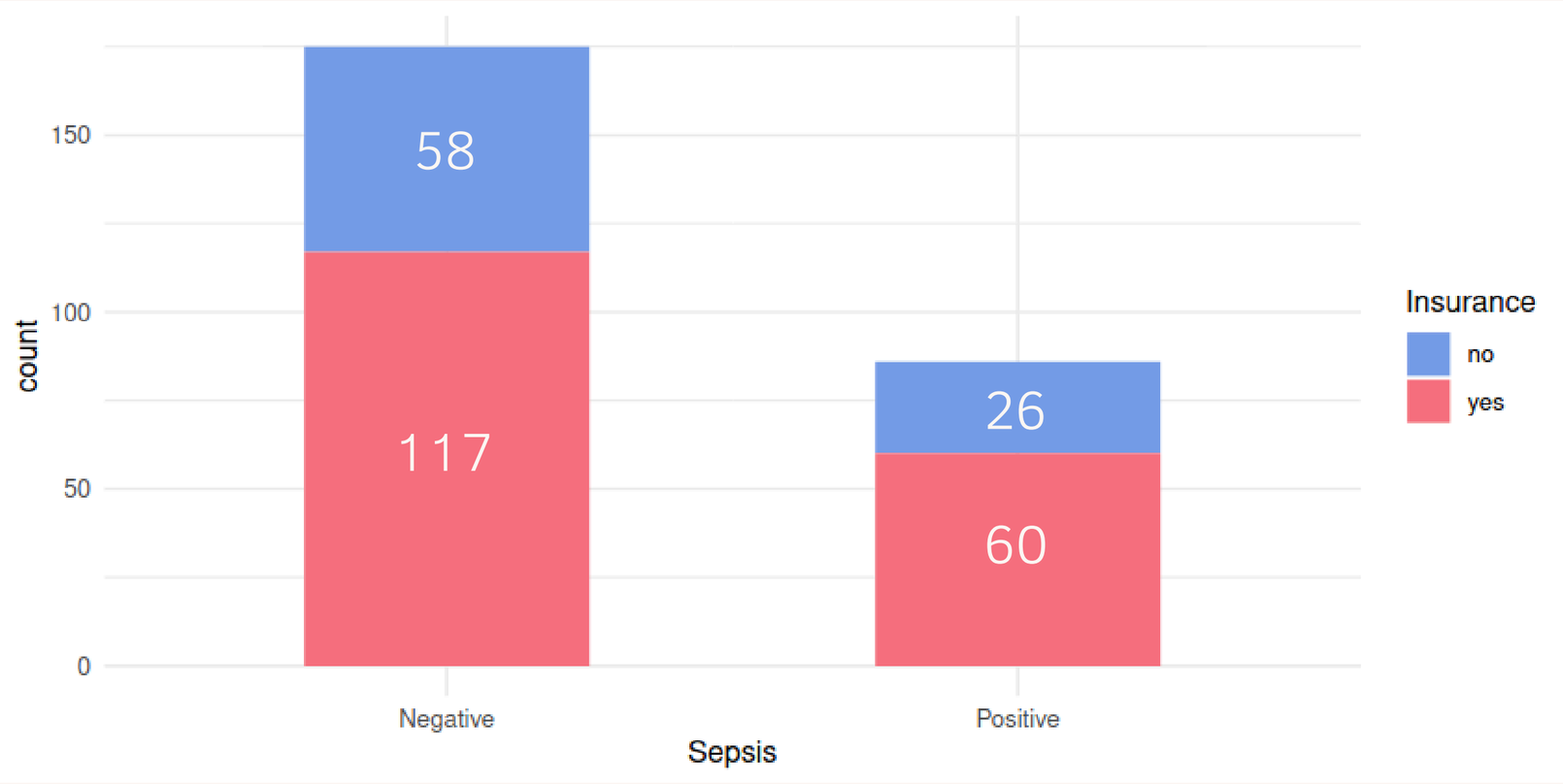
BD2: W = 0.9, p-value = 6e-14

Age: W = 0.9, p-value = 8e-15

NENHUMA DAS VARIÁVEIS APRESENTA DISTRIBUIÇÃO NORMAL

COMPARAÇÃO ENTRE DUAS MEDIDAS

Existe associação entre desfecho clínico e seguro saúde?



COMPARAÇÃO ENTRE DUAS MEDIDAS

Existe associação entre desfecho clínico e seguro saúde?

Teste exato de Fisher

> fisher.test(matriz_contingencia)

p-value = 0.7

95 percent confidence interval: 0.63 2.09

sample estimates: odds ratio 1.1

AO NÍVEL DE SIGNIFICÂNCIA DE 5%, NÃO EXISTE ESSA ASSOCIAÇÃO

COMPARAÇÃO ENTRE DUAS MEDIDAS

Comparação entre duas medianas com teste de Wilcoxon

Para cada uma das variáveis, verificamos se há diferença entre os grupos Sepse positivo e Sepse negativo

```
> wilcox.test(patients_positive$variavel, patients_negative$variavel, paired = FALSE)
```

PRG: W = 10130, p-value = 4e-06

PL: W = 12068, p-value = 2e-15

PR: W = 9851, p-value = 5e-05

SK: W = 10160, p-value = 4e-06

TS: W = 11812, p-value = 7e-14

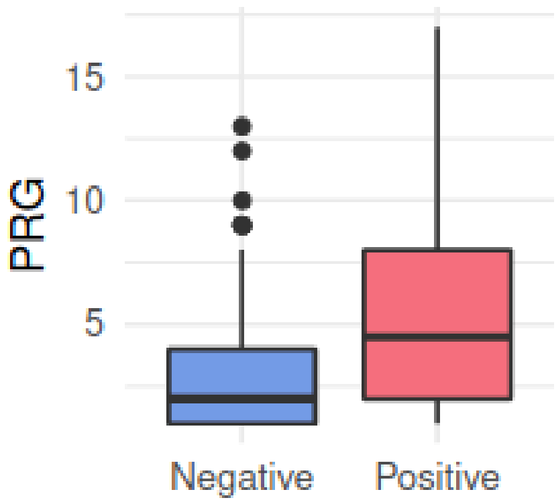
M11: W = 10148, p-value = 5e-06

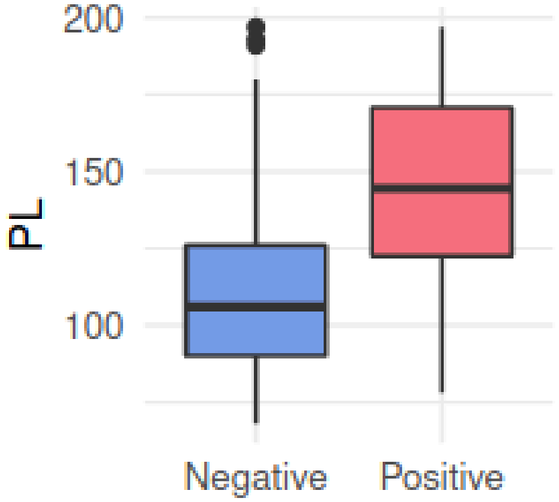
BD2: W = 9129, p-value = 0.005

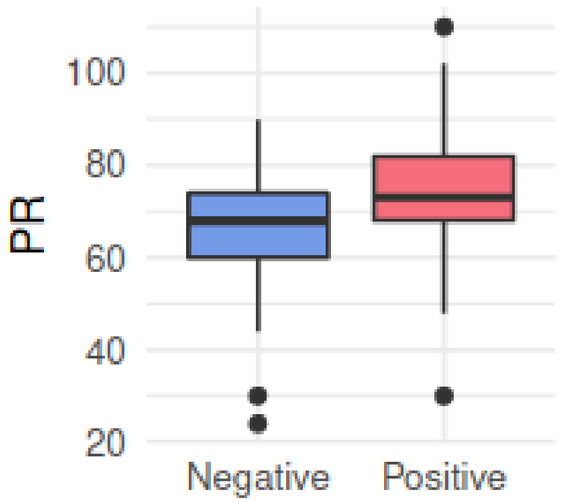
Age: W = 11612, p-value = 9e-13

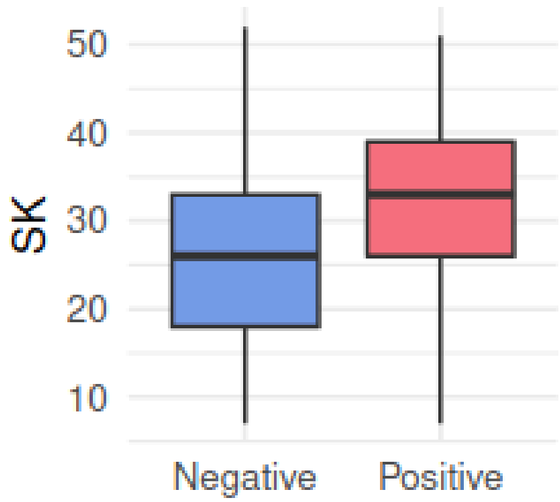
AO NÍVEL DE SIGNIFICÂNCIA DE 5%, TODAS AS VARIÁVEIS TESTADAS APRESENTAM MEDIANAS DIFERENTES ENTRE OS GRUPOS

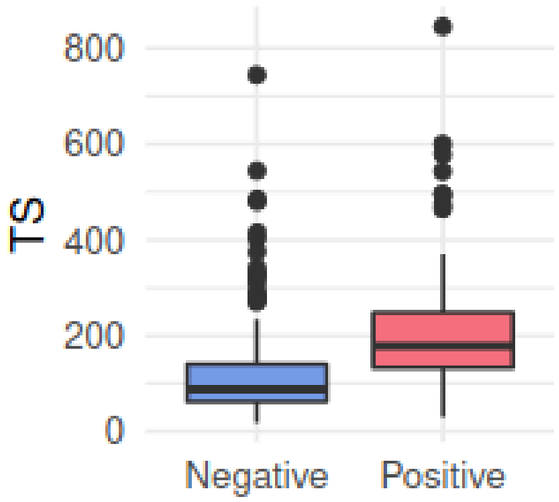
COMPARAÇÃO ENTRE DUAS MEDIDAS

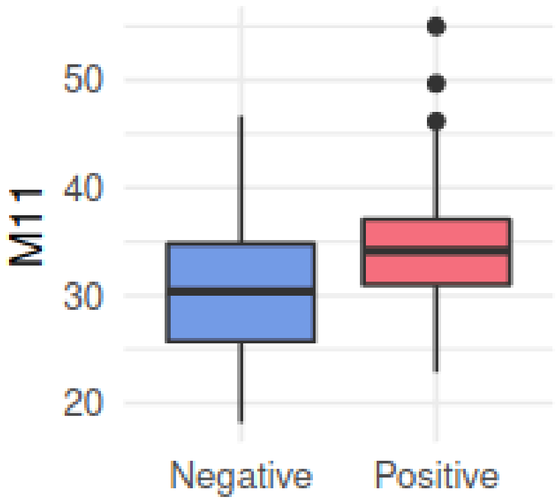


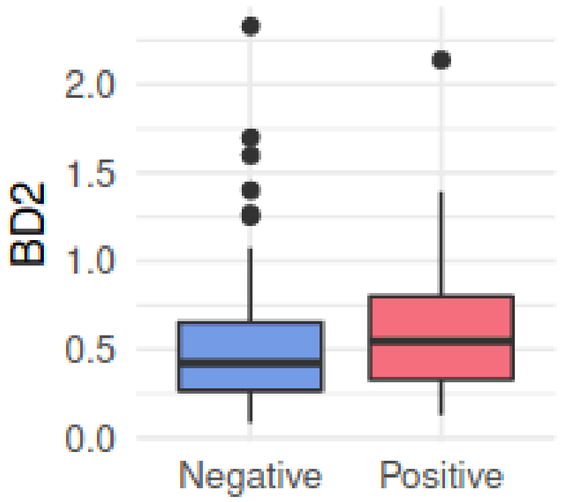


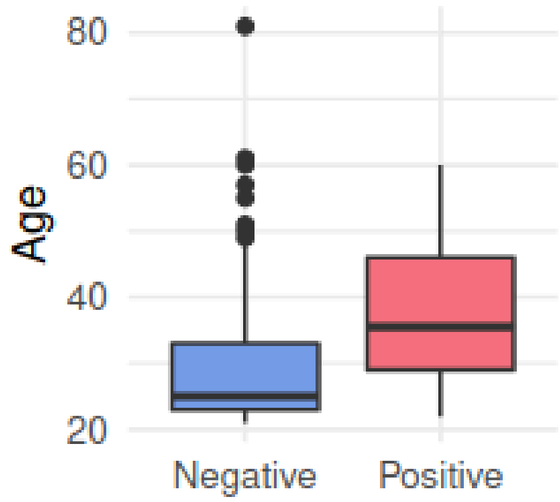










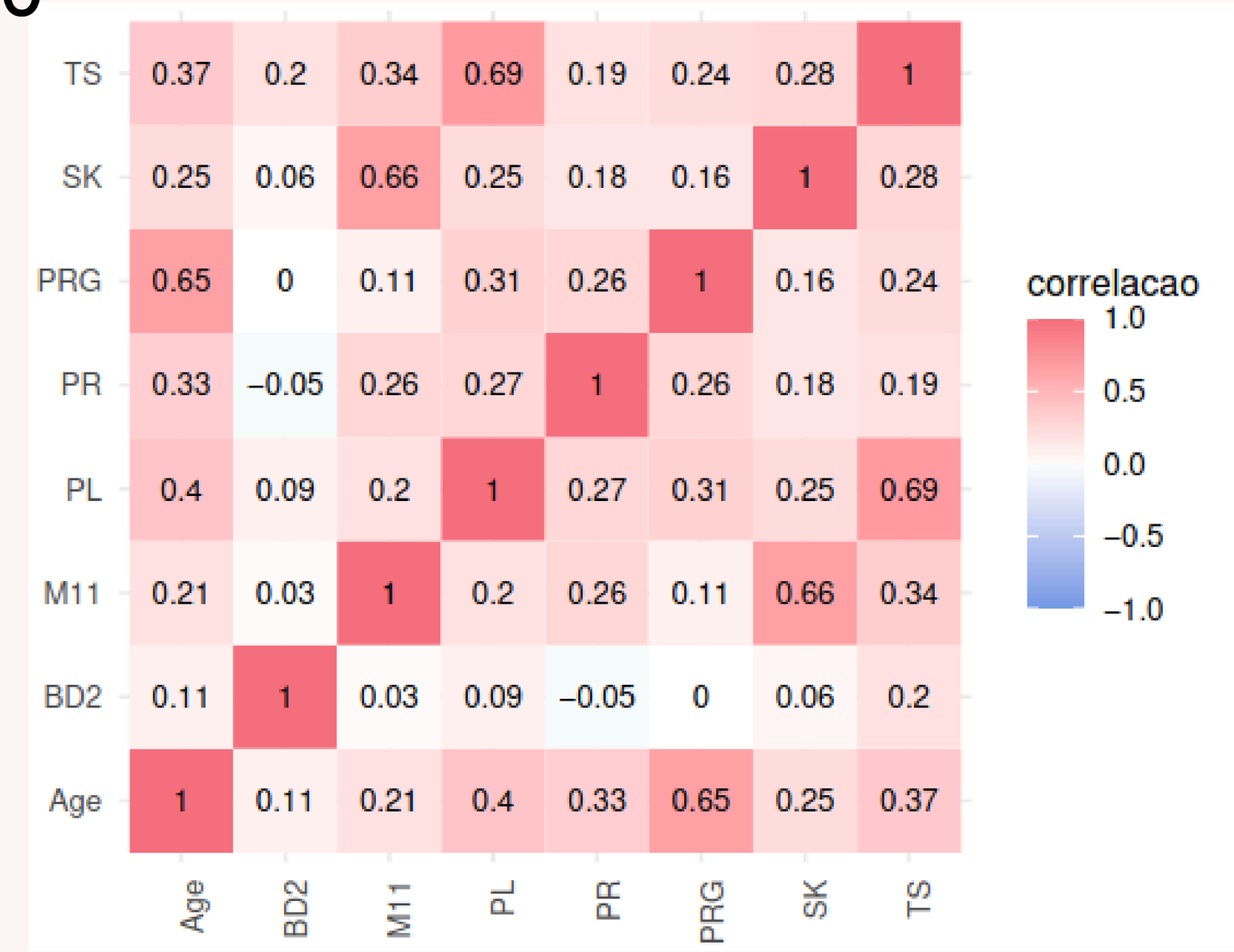


TESTE DE ASSOCIAÇÃO

Teste de correlação de Spearman

- Verificar se há associação entre duas variáveis

```
> cor(patients[,variavel],
        method = "spearman")
```



ASSOCIAÇÃO ENTRE DUAS VARIÁVEIS