

Modelando dados de incidência de câncer de próstata

Jailson Rodrigues de souza

Universidade Federal do Ceará, Fortaleza, 63100, Ceará, Brasil

Resumo The attendance system at Multimedia University has evolved from signing on papers to a QR code based attendance system, whereby students were to scan the QR code projected by a lecturer to register their attendance for the class. This paper examines the approach of a QR code based attendance system and proposes a method to exploit the oversight of it's implementation.

Keywords: Modelos de Regressão · Linear · Câncer

1 Introdução

Um grupo de pesquisadores de um determinado centro médico universitário está interessado em estudar a associação entre antígeno específico da próstata (PSA) e algumas medidas clínicas prognósticas em homens com câncer de próstata em estado avançado. Os dados foram coletados de 97 homens que estavam prestes a sofrer prostatectomias radicais. O conjunto de dados possui um número identificando o paciente e informações a respeito de 8 medidas clínicas.

2 Análise Descritiva

```
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(ggplot2)
library(data.table)
```

```

##
## Attaching package: 'data.table'
## The following objects are masked from 'package:dplyr':
##
##   between, first, last

library(rlang)

##
## Attaching package: 'rlang'
## The following object is masked from 'package:data.table':
##
##   :=

library(purrr)

##
## Attaching package: 'purrr'
## The following objects are masked from 'package:rlang':
##
##   %%, as_function, flatten, flatten_chr, flatten_dbl,
##   flatten_int, flatten_lgl, flatten_raw, invoke, list_along,
##   modify, prepend, splice
## The following object is masked from 'package:data.table':
##
##   transpose

library("cowplot")

##
## *****
## Note: As of version 1.0.0, cowplot does not change the
## default ggplot2 theme anymore. To recover the previous
## behavior, execute:
##   theme_set(theme_cowplot())
## *****

library(MASS)

##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##
##   select

library("kableExtra")

```

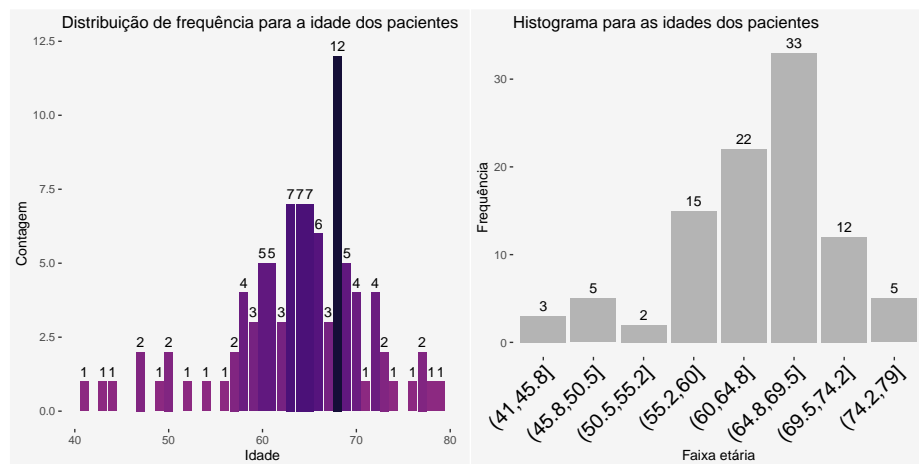
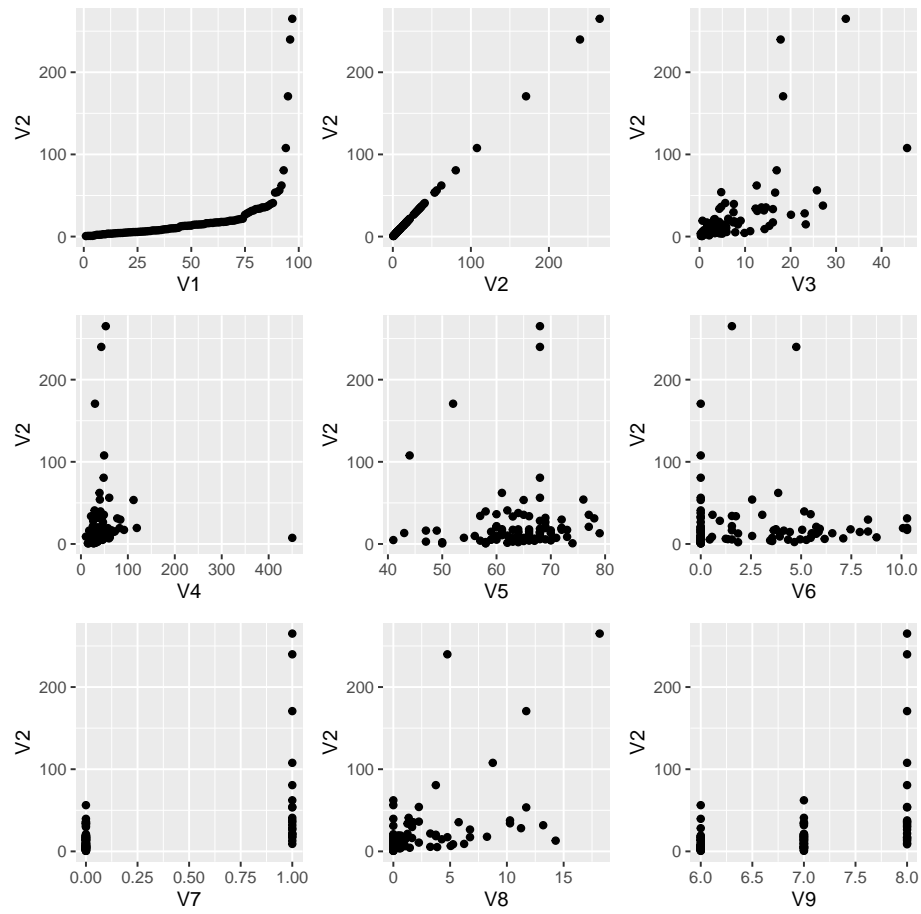
```
##
## Attaching package: 'kableExtra'
## The following object is masked from 'package:dplyr':
##
## group_rows
```

Tabela 1. Descrição das variáveis utilizadas no estudo

Número da variável	Nome da variável	Descrição
1	Número de identificação	1-97
2	Nível PSA	Nível sérico de antígeno prostático específico (mg/ml)
3	Volume câncer	Estimativa do volume do câncer (cc)
4	Peso	Peso da próstata (gm)
5	Idade	Idade do paciente (anos)
6	Hiperplasia prostática benigna	Quantidade de hiperplasia prostática benigna (cm ²)
7	Invasão da vesícula seminal	Presença ou ausência (1 se sim; 0 se não)
8	Penetração capsular	Grau de penetração capsular (cm)
9	Escore Gleason	Grau patologicamente determinado da doença (escores altos indicam pior prognóstico)

Fonte: U.S. Census, 2010

Vamos começar plotando as correlações marginais para buscar indícios que nos levem a encontrar fatores iniciais para a análise



3 Inferência e Modelagem

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
V1	1	57997.187058	57997.187058	63.6652953	0.0000000
V3	1	16228.755327	16228.755327	17.8148037	0.0000591
V4	1	40.078393	40.078393	0.0439953	0.8343472
V5	1	732.382105	732.382105	0.8039584	0.3723589
V6	1	1.865549	1.865549	0.0020479	0.9640079
V7	1	2615.345524	2615.345524	2.8709452	0.0937283
V8	1	1596.398534	1596.398534	1.7524158	0.1890019
V9	1	294.318419	294.318419	0.3230824	0.5712088
Residuals	88	80165.377963	910.970204	NA	NA

3.1 Selecção de Variáveis

4 Diagnóstico

5 Conclusões

Referências

1. Leslie Lamport, *L^AT_EX: a document preparation system*, Addison Wesley, Massachusetts, 2nd edition, 1994.