Projeto - Data Science Academy

André Campos da Silva

8 de Novembro, 2020

Projeto - Analise de Crédito

Fazer uma analise de credito em clientes novos, baseados em dados de clientes antigos, afim de determinar se esse novo cliente pode ou não ter carta de crédtio. ## Coletando os dados

```
# Carrego os pacotes necessários para o projeto
library('tidyverse')
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library("plyr")
```

```
## ------------------------------------------------------------------------------
```

```
## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:
## library(plyr); library(dplyr)
```

```
## ------------------------------------------------------------------------------
```

```
##
## Attaching package: 'plyr'
```

```
## The following objects are masked from 'package:dplyr':
##
##     arrange, count, desc, failwith, id, mutate, rename, summarise,
##     summarize
```

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

```
library('data.table')
```

```
##
## Attaching package: 'data.table'
```

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

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

```
library('corrplot')
```

```
## corrplot 0.84 loaded
```

```
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:dplyr':
##
##     combine
```

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

```
library('caTools')
library('DMwR')
```

```
## Loading required package: lattice
```

```
## Loading required package: grid
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
##
## Attaching package: 'DMwR'
```

```
## The following object is masked from 'package:plyr':
##
##     join
```

```
library('caret')
```

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

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

```
library('caTools')
```

```
# Carrego os dados de treino que será tratado e usado para a análise e treinamento.
df_credito_base <- read_csv('credit_dataset.csv')
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_double()
## )
## i Use `spec()` for the full column specifications.
```

```
head(df_credito_base)
```

```
## # A tibble: 6 x 21
##   credit.rating account.balance credit.duration~ previous.credit~ credit.purpose
##           <dbl>           <dbl>            <dbl>            <dbl>          <dbl>
## 1             1               1               18                3              2
## 2             1               1                9                3              4
## 3             1               2               12                2              4
## 4             1               1               12                3              4
## 5             1               1               12                3              4
## 6             1               1               10                3              4
## # ... with 16 more variables: credit.amount <dbl>, savings <dbl>,
## #   employment.duration <dbl>, installment.rate <dbl>, marital.status <dbl>,
## #   guarantor <dbl>, residence.duration <dbl>, current.assets <dbl>, age <dbl>,
## #   other.credits <dbl>, apartment.type <dbl>, bank.credits <dbl>,
## #   occupation <dbl>, dependents <dbl>, telephone <dbl>, foreign.worker <dbl>
```

```
# Faço uma verificação do formato dos dados.
glimpse(df_credito_base)
```

```
## Rows: 1,000
## Columns: 21
## $ credit.rating                  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ account.balance                <dbl> 1, 1, 2, 1, 1, 1, 1, 1, 3, 2, 1, 1, ...
## $ credit.duration.months         <dbl> 18, 9, 12, 12, 12, 10, 8, 6, 18, 24,...
## $ previous.credit.payment.status <dbl> 3, 3, 2, 3, 3, 3, 3, 3, 3, 2, 3, 3, ...
## $ credit.purpose                 <dbl> 2, 4, 4, 4, 4, 4, 4, 4, 3, 3, 4, 1, ...
## $ credit.amount                  <dbl> 1049, 2799, 841, 2122, 2171, 2241, 3...
## $ savings                        <dbl> 1, 1, 2, 1, 1, 1, 1, 1, 1, 3, 1, 2, ...
## $ employment.duration            <dbl> 1, 2, 3, 2, 2, 1, 3, 1, 1, 1, 2, 3, ...
## $ installment.rate               <dbl> 4, 2, 2, 3, 4, 1, 1, 2, 4, 1, 2, 1, ...
## $ marital.status                 <dbl> 1, 3, 1, 3, 3, 3, 3, 3, 1, 1, 3, 4, ...
## $ guarantor                      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ residence.duration             <dbl> 4, 2, 4, 2, 4, 3, 4, 4, 4, 4, 2, 4, ...
## $ current.assets                 <dbl> 2, 1, 1, 1, 2, 1, 1, 1, 3, 4, 1, 3, ...
## $ age                            <dbl> 21, 36, 23, 39, 38, 48, 39, 40, 65, ...
## $ other.credits                  <dbl> 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, ...
## $ apartment.type                 <dbl> 1, 1, 1, 1, 2, 1, 2, 2, 2, 1, 1, 1, ...
## $ bank.credits                   <dbl> 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2, 2, ...
## $ occupation                     <dbl> 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 3, 3, ...
## $ dependents                     <dbl> 1, 2, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, ...
## $ telephone                      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ foreign.worker                 <dbl> 1, 1, 1, 2, 2, 2, 2, 2, 1, 1, 1, 1, ...
```

Tratamento dos dados

```
# Os dados estão quase todos na mesmas escala apenas 3 que estão fora, esses nos irei
# normaliza-los, os demais são fatores e então colocarei como fator
# sobretudo a variavel target credit.rating.
scale(df_credito_base$age, center = T, scale = T)
```

```
##               [,1]
##    [1,] -1.28093214
##    [2,]  0.04034293
##    [3,] -1.10476213
##    [4,]  0.30459795
##    [5,]  0.21651294
##    [6,]  1.09736299
##    [7,]  0.30459795
##    [8,]  0.39268295
##    [9,]  2.59480806
##   [10,] -1.10476213
##   [11,]  0.04034293
##   [12,] -1.01667712
##   [13,] -0.40008209
##   [14,] -0.40008209
##   [15,] -1.10476213
##   [16,]  0.74502297
##   [17,]  0.39268295
##   [18,] -0.92859212
##   [19,]  0.04034293
##   [20,]  0.30459795
##   [21,]  0.12842794
##   [22,]  1.18544799
##   [23,] -0.22391208
##   [24,] -0.84050711
##   [25,]  0.74502297
##   [26,]  1.36161800
##   [27,] -1.01667712
##   [28,]  0.12842794
##   [29,] -0.57625210
##   [30,]  1.80204302
##   [31,]  1.00927798
##   [32,]  0.12842794
##   [33,] -0.13582708
##   [34,] -0.66433710
##   [35,]  0.04034293
##   [36,]  0.30459795
##   [37,]  0.21651294
##   [38,] -0.84050711
##   [39,] -0.40008209
##   [40,]  0.48076795
##   [41,] -1.10476213
##   [42,]  1.97821303
##   [43,]  0.04034293
##   [44,] -0.13582708
##   [45,]  2.24246805
##   [46,] -1.10476213
##   [47,] -0.92859212
##   [48,] -0.48816710
##   [49,] -0.22391208
##   [50,] -0.84050711
##   [51,]  1.18544799
##   [52,] -1.10476213
```

```
##   [53,] -0.48816710
##   [54,]  1.18544799
##   [55,]  0.48076795
##   [56,]  1.97821303
##   [57,]  2.41863805
##   [58,] -0.75242211
##   [59,]  0.48076795
##   [60,]  0.04034293
##   [61,]  0.74502297
##   [62,] -0.57625210
##   [63,]  1.00927798
##   [64,]  0.83310797
##   [65,]  0.65693796
##   [66,] -0.75242211
##   [67,]  1.44970300
##   [68,] -0.66433710
##   [69,] -1.19284713
##   [70,]  2.15438304
##   [71,] -0.31199709
##   [72,]  1.97821303
##   [73,]  0.30459795
##   [74,] -0.40008209
##   [75,]  0.12842794
##   [76,] -1.01667712
##   [77,]  0.04034293
##   [78,]  2.15438304
##   [79,] -0.48816710
##   [80,] -0.04774207
##   [81,]  0.12842794
##   [82,]  1.36161800
##   [83,] -0.48816710
##   [84,]  0.04034293
##   [85,]  2.41863805
##   [86,] -0.40008209
##   [87,] -0.04774207
##   [88,] -0.13582708
##   [89,]  0.30459795
##   [90,]  0.56885296
##   [91,] -0.75242211
##   [92,] -0.92859212
##   [93,] -0.66433710
##   [94,]  2.06629804
##   [95,]  0.48076795
##   [96,] -0.31199709
##   [97,] -1.10476213
##   [98,] -0.48816710
##   [99,] -0.92859212
##  [100,]  0.48076795
##  [101,]  0.39268295
##  [102,] -1.19284713
##  [103,] -0.13582708
##  [104,]  1.62587301
##  [105,] -0.22391208
##  [106,] -0.40008209
```

```
##  [107,] -0.57625210
##  [108,]  0.30459795
##  [109,]  0.21651294
##  [110,]  2.24246805
##  [111,] -0.57625210
##  [112,]  1.18544799
##  [113,]  0.04034293
##  [114,] -0.84050711
##  [115,] -0.04774207
##  [116,] -0.66433710
##  [117,] -0.04774207
##  [118,] -0.66433710
##  [119,]  1.44970300
##  [120,] -0.31199709
##  [121,] -0.48816710
##  [122,] -0.04774207
##  [123,] -0.48816710
##  [124,]  0.65693796
##  [125,]  1.00927798
##  [126,]  1.00927798
##  [127,] -1.01667712
##  [128,] -0.31199709
##  [129,] -0.48816710
##  [130,]  0.04034293
##  [131,] -0.84050711
##  [132,]  0.65693796
##  [133,]  0.48076795
##  [134,] -0.84050711
##  [135,]  2.50672306
##  [136,] -0.22391208
##  [137,]  1.62587301
##  [138,] -0.31199709
##  [139,]  0.30459795
##  [140,] -0.40008209
##  [141,]  0.83310797
##  [142,]  2.50672306
##  [143,]  0.56885296
##  [144,] -0.40008209
##  [145,] -0.92859212
##  [146,]  0.92119298
##  [147,]  1.18544799
##  [148,]  1.18544799
##  [149,]  0.04034293
##  [150,]  1.09736299
##  [151,]  0.56885296
##  [152,] -0.57625210
##  [153,] -0.04774207
##  [154,]  0.21651294
##  [155,] -0.84050711
##  [156,] -0.04774207
##  [157,] -1.10476213
##  [158,]  0.21651294
##  [159,]  2.06629804
##  [160,] -0.66433710
```

```
##  [161,] -0.04774207
##  [162,]  0.04034293
##  [163,] -0.92859212
##  [164,] -0.66433710
##  [165,] -1.10476213
##  [166,]  1.09736299
##  [167,] -0.92859212
##  [168,]  3.38757310
##  [169,] -1.01667712
##  [170,] -0.66433710
##  [171,] -1.01667712
##  [172,] -1.01667712
##  [173,] -1.10476213
##  [174,] -0.31199709
##  [175,] -0.40008209
##  [176,]  0.04034293
##  [177,] -0.75242211
##  [178,]  0.74502297
##  [179,] -0.13582708
##  [180,]  0.12842794
##  [181,] -1.10476213
##  [182,] -0.48816710
##  [183,] -0.92859212
##  [184,] -0.22391208
##  [185,] -0.75242211
##  [186,] -0.40008209
##  [187,]  1.27353300
##  [188,] -0.40008209
##  [189,] -0.31199709
##  [190,] -0.66433710
##  [191,] -0.57625210
##  [192,] -0.75242211
##  [193,] -1.01667712
##  [194,]  1.62587301
##  [195,] -0.92859212
##  [196,] -0.04774207
##  [197,]  0.83310797
##  [198,] -0.66433710
##  [199,] -0.04774207
##  [200,] -0.75242211
##  [201,] -0.48816710
##  [202,]  0.39268295
##  [203,]  1.18544799
##  [204,] -0.48816710
##  [205,] -1.01667712
##  [206,] -0.22391208
##  [207,] -1.36901714
##  [208,] -1.10476213
##  [209,]  0.04034293
##  [210,] -0.04774207
##  [211,] -0.57625210
##  [212,] -1.19284713
##  [213,] -1.19284713
##  [214,]  3.38757310
```

```
##  [215,] -0.04774207
##  [216,]  0.12842794
##  [217,] -0.40008209
##  [218,] -0.75242211
##  [219,] -0.04774207
##  [220,]  0.83310797
##  [221,]  0.04034293
##  [222,]  1.18544799
##  [223,] -0.04774207
##  [224,] -0.75242211
##  [225,] -0.66433710
##  [226,]  0.65693796
##  [227,] -0.04774207
##  [228,] -0.75242211
##  [229,]  0.04034293
##  [230,] -0.48816710
##  [231,]  0.04034293
##  [232,]  0.30459795
##  [233,]  0.83310797
##  [234,]  0.39268295
##  [235,] -0.92859212
##  [236,]  1.71395802
##  [237,]  1.27353300
##  [238,]  2.50672306
##  [239,] -1.10476213
##  [240,] -1.01667712
##  [241,]  1.00927798
##  [242,]  0.56885296
##  [243,] -1.01667712
##  [244,] -0.31199709
##  [245,]  0.83310797
##  [246,] -1.10476213
##  [247,] -1.01667712
##  [248,] -0.04774207
##  [249,] -0.92859212
##  [250,] -0.84050711
##  [251,]  0.65693796
##  [252,]  0.74502297
##  [253,]  0.21651294
##  [254,] -0.66433710
##  [255,]  0.21651294
##  [256,] -0.84050711
##  [257,] -0.75242211
##  [258,] -1.10476213
##  [259,]  1.44970300
##  [260,]  3.38757310
##  [261,] -0.75242211
##  [262,] -0.04774207
##  [263,] -1.10476213
##  [264,] -1.36901714
##  [265,] -0.13582708
##  [266,]  0.39268295
##  [267,]  0.12842794
##  [268,]  0.04034293
```

```
##  [269,] -1.19284713
##  [270,] -0.13582708
##  [271,]  1.27353300
##  [272,]  0.39268295
##  [273,] -1.10476213
##  [274,]  1.53778801
##  [275,] -0.04774207
##  [276,] -0.04774207
##  [277,]  1.62587301
##  [278,] -0.75242211
##  [279,] -0.66433710
##  [280,]  1.18544799
##  [281,] -0.57625210
##  [282,]  0.92119298
##  [283,] -1.36901714
##  [284,]  1.00927798
##  [285,] -0.57625210
##  [286,]  1.62587301
##  [287,]  1.44970300
##  [288,] -0.31199709
##  [289,]  1.36161800
##  [290,] -0.22391208
##  [291,]  0.12842794
##  [292,] -1.36901714
##  [293,] -0.13582708
##  [294,] -0.40008209
##  [295,] -0.48816710
##  [296,] -0.48816710
##  [297,] -1.19284713
##  [298,] -0.13582708
##  [299,] -1.45710214
##  [300,] -0.13582708
##  [301,]  0.83310797
##  [302,] -0.92859212
##  [303,] -0.84050711
##  [304,] -0.22391208
##  [305,] -0.66433710
##  [306,]  0.48076795
##  [307,]  0.04034293
##  [308,] -0.84050711
##  [309,] -0.66433710
##  [310,]  0.30459795
##  [311,]  2.59480806
##  [312,]  0.21651294
##  [313,]  0.48076795
##  [314,] -0.84050711
##  [315,]  0.48076795
##  [316,]  0.83310797
##  [317,]  0.39268295
##  [318,] -1.01667712
##  [319,] -0.40008209
##  [320,] -0.04774207
##  [321,]  0.39268295
##  [322,] -0.66433710
```

```
##  [323,] -0.66433710
##  [324,] -0.75242211
##  [325,] -0.40008209
##  [326,]  0.12842794
##  [327,] -0.04774207
##  [328,]  0.92119298
##  [329,] -0.31199709
##  [330,] -0.22391208
##  [331,] -0.22391208
##  [332,] -0.84050711
##  [333,] -1.01667712
##  [334,] -0.92859212
##  [335,]  1.62587301
##  [336,] -1.01667712
##  [337,]  1.09736299
##  [338,]  1.18544799
##  [339,] -0.31199709
##  [340,]  1.71395802
##  [341,]  0.39268295
##  [342,] -1.19284713
##  [343,]  0.21651294
##  [344,] -0.84050711
##  [345,]  1.36161800
##  [346,]  0.21651294
##  [347,]  0.92119298
##  [348,]  2.41863805
##  [349,] -0.48816710
##  [350,] -0.31199709
##  [351,] -0.84050711
##  [352,] -0.13582708
##  [353,]  0.92119298
##  [354,] -0.57625210
##  [355,]  0.56885296
##  [356,]  0.56885296
##  [357,]  2.41863805
##  [358,] -0.31199709
##  [359,] -1.10476213
##  [360,]  2.06629804
##  [361,]  2.24246805
##  [362,] -1.19284713
##  [363,] -0.75242211
##  [364,] -0.84050711
##  [365,]  0.12842794
##  [366,] -0.13582708
##  [367,]  1.89012803
##  [368,] -0.40008209
##  [369,] -0.66433710
##  [370,] -1.01667712
##  [371,] -1.28093214
##  [372,] -0.13582708
##  [373,] -0.57625210
##  [374,] -0.75242211
##  [375,] -0.13582708
##  [376,] -0.75242211
```

```
##  [377,] -0.66433710
##  [378,] -0.48816710
##  [379,] -1.19284713
##  [380,]  0.04034293
##  [381,] -1.19284713
##  [382,] -0.84050711
##  [383,]  2.41863805
##  [384,] -1.19284713
##  [385,]  0.04034293
##  [386,] -1.01667712
##  [387,] -0.48816710
##  [388,]  0.65693796
##  [389,]  0.12842794
##  [390,] -0.84050711
##  [391,] -0.22391208
##  [392,]  2.68289307
##  [393,]  1.27353300
##  [394,] -0.04774207
##  [395,] -0.92859212
##  [396,] -0.92859212
##  [397,] -0.22391208
##  [398,] -0.48816710
##  [399,]  0.48076795
##  [400,] -0.31199709
##  [401,] -0.31199709
##  [402,]  0.30459795
##  [403,] -1.10476213
##  [404,] -0.84050711
##  [405,] -0.31199709
##  [406,]  0.39268295
##  [407,] -0.48816710
##  [408,] -0.75242211
##  [409,]  0.04034293
##  [410,] -0.57625210
##  [411,]  2.50672306
##  [412,] -0.31199709
##  [413,] -0.04774207
##  [414,]  0.74502297
##  [415,]  0.56885296
##  [416,] -0.57625210
##  [417,]  1.89012803
##  [418,] -0.75242211
##  [419,] -0.57625210
##  [420,]  0.74502297
##  [421,] -0.66433710
##  [422,]  0.39268295
##  [423,]  2.85906308
##  [424,]  1.44970300
##  [425,]  0.92119298
##  [426,]  1.44970300
##  [427,] -0.31199709
##  [428,] -0.13582708
##  [429,] -1.10476213
##  [430,] -0.22391208
```

```
##  [431,] -0.84050711
##  [432,]  1.71395802
##  [433,] -0.75242211
##  [434,] -0.66433710
##  [435,]  1.89012803
##  [436,]  0.04034293
##  [437,] -0.40008209
##  [438,] -1.01667712
##  [439,]  0.65693796
##  [440,]  0.39268295
##  [441,] -1.10476213
##  [442,] -0.84050711
##  [443,] -0.48816710
##  [444,]  0.39268295
##  [445,]  1.80204302
##  [446,]  1.27353300
##  [447,]  0.04034293
##  [448,] -1.28093214
##  [449,] -0.31199709
##  [450,]  2.68289307
##  [451,]  2.68289307
##  [452,]  0.30459795
##  [453,]  3.03523309
##  [454,] -1.01667712
##  [455,] -1.19284713
##  [456,] -0.75242211
##  [457,] -0.57625210
##  [458,] -0.04774207
##  [459,] -0.57625210
##  [460,] -0.84050711
##  [461,]  2.15438304
##  [462,]  0.39268295
##  [463,] -1.10476213
##  [464,]  0.83310797
##  [465,] -1.01667712
##  [466,]  1.71395802
##  [467,] -0.48816710
##  [468,] -0.04774207
##  [469,] -1.01667712
##  [470,]  1.00927798
##  [471,]  1.00927798
##  [472,] -0.92859212
##  [473,]  0.12842794
##  [474,] -0.22391208
##  [475,] -0.40008209
##  [476,] -0.75242211
##  [477,]  1.00927798
##  [478,] -0.92859212
##  [479,]  1.89012803
##  [480,]  2.77097807
##  [481,]  1.09736299
##  [482,]  0.39268295
##  [483,] -0.31199709
##  [484,]  1.18544799
```

```
##  [485,] -0.04774207
##  [486,] -0.75242211
##  [487,] -0.04774207
##  [488,]  0.12842794
##  [489,] -0.84050711
##  [490,]  0.83310797
##  [491,] -0.22391208
##  [492,]  0.56885296
##  [493,] -0.40008209
##  [494,]  0.92119298
##  [495,]  1.62587301
##  [496,] -0.84050711
##  [497,] -0.57625210
##  [498,]  0.65693796
##  [499,]  0.21651294
##  [500,] -0.04774207
##  [501,] -0.66433710
##  [502,] -0.75242211
##  [503,]  1.36161800
##  [504,] -0.22391208
##  [505,]  0.21651294
##  [506,] -0.48816710
##  [507,]  0.21651294
##  [508,]  0.92119298
##  [509,] -0.22391208
##  [510,]  0.65693796
##  [511,] -1.01667712
##  [512,]  1.80204302
##  [513,]  0.12842794
##  [514,] -1.19284713
##  [515,] -1.36901714
##  [516,] -0.57625210
##  [517,]  0.04034293
##  [518,]  0.12842794
##  [519,]  0.74502297
##  [520,] -1.10476213
##  [521,] -0.66433710
##  [522,] -0.84050711
##  [523,] -0.57625210
##  [524,] -1.28093214
##  [525,] -0.13582708
##  [526,] -0.40008209
##  [527,]  0.39268295
##  [528,] -0.75242211
##  [529,] -0.31199709
##  [530,] -0.92859212
##  [531,]  0.56885296
##  [532,]  0.65693796
##  [533,]  0.74502297
##  [534,] -0.31199709
##  [535,] -0.75242211
##  [536,] -1.36901714
##  [537,]  0.12842794
##  [538,]  0.56885296
```

```
##  [539,]  0.92119298
##  [540,] -0.22391208
##  [541,] -1.19284713
##  [542,]  0.48076795
##  [543,]  0.12842794
##  [544,] -0.75242211
##  [545,]  0.56885296
##  [546,] -0.84050711
##  [547,] -0.84050711
##  [548,] -0.75242211
##  [549,] -1.19284713
##  [550,]  0.12842794
##  [551,] -0.48816710
##  [552,]  1.18544799
##  [553,] -0.75242211
##  [554,]  0.39268295
##  [555,] -1.36901714
##  [556,] -0.84050711
##  [557,] -0.48816710
##  [558,] -0.22391208
##  [559,] -0.84050711
##  [560,] -0.04774207
##  [561,]  0.74502297
##  [562,] -0.75242211
##  [563,] -1.01667712
##  [564,]  1.00927798
##  [565,] -0.84050711
##  [566,] -1.10476213
##  [567,] -0.40008209
##  [568,]  1.44970300
##  [569,]  0.65693796
##  [570,]  0.12842794
##  [571,] -0.84050711
##  [572,] -0.04774207
##  [573,] -0.48816710
##  [574,] -0.13582708
##  [575,] -0.04774207
##  [576,]  1.09736299
##  [577,]  0.56885296
##  [578,]  0.39268295
##  [579,] -1.10476213
##  [580,] -0.48816710
##  [581,] -1.01667712
##  [582,] -0.22391208
##  [583,] -0.40008209
##  [584,] -1.19284713
##  [585,]  1.53778801
##  [586,] -0.13582708
##  [587,]  0.04034293
##  [588,]  0.12842794
##  [589,] -0.92859212
##  [590,] -0.13582708
##  [591,] -0.75242211
##  [592,]  1.36161800
```

```
##  [593,] -1.10476213
##  [594,] -0.57625210
##  [595,] -0.31199709
##  [596,] -0.66433710
##  [597,]  0.21651294
##  [598,]  2.59480806
##  [599,]  0.30459795
##  [600,] -0.57625210
##  [601,] -0.84050711
##  [602,] -0.48816710
##  [603,]  3.47565811
##  [604,] -1.28093214
##  [605,]  0.48076795
##  [606,] -0.92859212
##  [607,] -0.04774207
##  [608,]  0.12842794
##  [609,] -0.84050711
##  [610,] -0.48816710
##  [611,] -1.10476213
##  [612,]  0.56885296
##  [613,] -0.48816710
##  [614,]  0.92119298
##  [615,] -0.57625210
##  [616,] -0.57625210
##  [617,]  0.21651294
##  [618,] -0.48816710
##  [619,] -0.40008209
##  [620,]  1.00927798
##  [621,]  1.09736299
##  [622,] -0.75242211
##  [623,]  2.68289307
##  [624,] -0.57625210
##  [625,]  3.47565811
##  [626,]  0.04034293
##  [627,]  0.21651294
##  [628,]  0.92119298
##  [629,]  0.04034293
##  [630,]  1.71395802
##  [631,] -0.92859212
##  [632,]  0.83310797
##  [633,] -0.66433710
##  [634,] -0.48816710
##  [635,]  0.30459795
##  [636,] -1.19284713
##  [637,]  2.15438304
##  [638,] -0.66433710
##  [639,] -1.01667712
##  [640,]  2.41863805
##  [641,]  0.92119298
##  [642,] -0.22391208
##  [643,]  0.74502297
##  [644,]  1.97821303
##  [645,]  1.62587301
##  [646,]  0.65693796
```

```
##  [647,]  1.44970300
##  [648,] -0.31199709
##  [649,]  0.48076795
##  [650,] -0.13582708
##  [651,]  1.09736299
##  [652,]  0.21651294
##  [653,]  0.21651294
##  [654,]  1.53778801
##  [655,]  0.21651294
##  [656,]  0.56885296
##  [657,]  0.21651294
##  [658,]  2.50672306
##  [659,] -0.04774207
##  [660,] -0.84050711
##  [661,] -0.22391208
##  [662,] -0.57625210
##  [663,] -1.36901714
##  [664,] -0.66433710
##  [665,]  1.00927798
##  [666,]  1.18544799
##  [667,]  0.83310797
##  [668,] -1.10476213
##  [669,] -0.92859212
##  [670,] -0.13582708
##  [671,] -1.10476213
##  [672,] -0.31199709
##  [673,] -0.84050711
##  [674,]  0.74502297
##  [675,] -0.84050711
##  [676,] -0.66433710
##  [677,] -0.92859212
##  [678,]  1.09736299
##  [679,] -0.04774207
##  [680,] -0.66433710
##  [681,] -1.10476213
##  [682,]  0.39268295
##  [683,]  0.04034293
##  [684,]  1.27353300
##  [685,] -1.01667712
##  [686,]  0.39268295
##  [687,] -0.40008209
##  [688,] -0.57625210
##  [689,]  0.83310797
##  [690,] -0.31199709
##  [691,] -1.28093214
##  [692,] -0.75242211
##  [693,] -1.19284713
##  [694,]  0.92119298
##  [695,] -1.01667712
##  [696,] -1.28093214
##  [697,]  0.04034293
##  [698,] -0.66433710
##  [699,] -0.75242211
##  [700,]  1.89012803
```

```
##  [701,]  1.00927798
##  [702,] -0.13582708
##  [703,] -1.10476213
##  [704,]  1.89012803
##  [705,]  2.77097807
##  [706,]  0.04034293
##  [707,] -1.36901714
##  [708,]  0.04034293
##  [709,]  0.12842794
##  [710,] -0.40008209
##  [711,]  2.33055305
##  [712,]  2.59480806
##  [713,]  0.04034293
##  [714,]  0.48076795
##  [715,] -0.84050711
##  [716,] -0.75242211
##  [717,]  1.27353300
##  [718,]  1.62587301
##  [719,] -0.75242211
##  [720,] -1.01667712
##  [721,]  1.27353300
##  [722,]  0.39268295
##  [723,] -0.66433710
##  [724,] -0.40008209
##  [725,] -0.75242211
##  [726,] -1.28093214
##  [727,] -0.40008209
##  [728,]  0.30459795
##  [729,] -0.48816710
##  [730,] -0.75242211
##  [731,]  1.71395802
##  [732,] -0.13582708
##  [733,] -1.10476213
##  [734,] -1.10476213
##  [735,]  0.56885296
##  [736,]  2.33055305
##  [737,]  0.12842794
##  [738,]  0.92119298
##  [739,]  0.21651294
##  [740,] -0.40008209
##  [741,] -0.40008209
##  [742,]  2.24246805
##  [743,] -1.36901714
##  [744,]  2.41863805
##  [745,] -1.01667712
##  [746,] -1.28093214
##  [747,]  0.04034293
##  [748,]  0.74502297
##  [749,]  2.77097807
##  [750,]  1.36161800
##  [751,] -0.48816710
##  [752,] -0.66433710
##  [753,] -0.31199709
##  [754,] -0.66433710
```

```
##  [755,] -1.10476213
##  [756,] -0.22391208
##  [757,]  2.15438304
##  [758,]  1.09736299
##  [759,] -1.01667712
##  [760,] -0.75242211
##  [761,] -0.66433710
##  [762,]  2.59480806
##  [763,] -0.57625210
##  [764,]  0.56885296
##  [765,]  1.00927798
##  [766,] -1.19284713
##  [767,] -0.92859212
##  [768,] -0.31199709
##  [769,] -0.57625210
##  [770,] -1.01667712
##  [771,] -0.48816710
##  [772,] -0.75242211
##  [773,]  1.00927798
##  [774,] -1.01667712
##  [775,] -1.01667712
##  [776,] -0.13582708
##  [777,]  0.39268295
##  [778,]  0.74502297
##  [779,] -0.22391208
##  [780,]  0.39268295
##  [781,]  1.18544799
##  [782,]  1.09736299
##  [783,]  1.89012803
##  [784,] -1.01667712
##  [785,] -0.57625210
##  [786,] -0.04774207
##  [787,] -0.31199709
##  [788,]  2.24246805
##  [789,]  0.04034293
##  [790,] -0.92859212
##  [791,] -0.84050711
##  [792,] -0.92859212
##  [793,]  1.71395802
##  [794,]  1.53778801
##  [795,]  2.68289307
##  [796,]  2.24246805
##  [797,] -0.13582708
##  [798,] -0.22391208
##  [799,] -0.57625210
##  [800,]  0.21651294
##  [801,] -0.04774207
##  [802,] -0.31199709
##  [803,] -0.92859212
##  [804,] -0.22391208
##  [805,] -1.10476213
##  [806,] -1.01667712
##  [807,] -0.66433710
##  [808,]  2.41863805
```

```
##  [809,]  0.30459795
##  [810,] -0.04774207
##  [811,] -1.10476213
##  [812,]  0.04034293
##  [813,] -1.10476213
##  [814,] -1.01667712
##  [815,]  3.38757310
##  [816,] -1.19284713
##  [817,]  1.09736299
##  [818,] -0.84050711
##  [819,] -1.01667712
##  [820,] -1.45710214
##  [821,] -0.75242211
##  [822,] -0.57625210
##  [823,] -0.31199709
##  [824,] -0.92859212
##  [825,] -0.57625210
##  [826,]  1.00927798
##  [827,] -0.13582708
##  [828,]  0.65693796
##  [829,] -0.48816710
##  [830,] -1.19284713
##  [831,]  0.56885296
##  [832,] -0.84050711
##  [833,] -0.75242211
##  [834,] -0.84050711
##  [835,] -0.31199709
##  [836,] -1.10476213
##  [837,]  0.04034293
##  [838,]  0.65693796
##  [839,]  0.30459795
##  [840,] -0.84050711
##  [841,]  0.39268295
##  [842,] -0.40008209
##  [843,] -0.57625210
##  [844,] -0.57625210
##  [845,]  0.30459795
##  [846,] -1.10476213
##  [847,] -0.48816710
##  [848,]  2.15438304
##  [849,]  1.27353300
##  [850,]  1.00927798
##  [851,] -1.28093214
##  [852,] -1.01667712
##  [853,] -0.40008209
##  [854,]  0.12842794
##  [855,]  0.65693796
##  [856,]  0.92119298
##  [857,] -0.48816710
##  [858,] -0.40008209
##  [859,] -0.13582708
##  [860,] -0.22391208
##  [861,]  0.21651294
##  [862,] -1.01667712
```

```
##  [863,] -1.10476213
##  [864,] -0.40008209
##  [865,] -0.75242211
##  [866,]  0.30459795
##  [867,] -0.92859212
##  [868,] -0.66433710
##  [869,] -0.75242211
##  [870,] -0.92859212
##  [871,]  1.62587301
##  [872,] -0.92859212
##  [873,] -1.19284713
##  [874,] -0.66433710
##  [875,] -0.13582708
##  [876,] -0.66433710
##  [877,] -0.66433710
##  [878,] -0.40008209
##  [879,] -1.19284713
##  [880,]  0.39268295
##  [881,] -0.92859212
##  [882,] -0.13582708
##  [883,]  0.56885296
##  [884,] -0.92859212
##  [885,] -0.66433710
##  [886,] -1.01667712
##  [887,] -0.75242211
##  [888,] -0.92859212
##  [889,] -0.04774207
##  [890,] -0.57625210
##  [891,] -0.75242211
##  [892,] -1.01667712
##  [893,] -0.92859212
##  [894,] -0.22391208
##  [895,] -0.13582708
##  [896,]  0.21651294
##  [897,]  0.48076795
##  [898,] -1.28093214
##  [899,]  0.04034293
##  [900,] -0.75242211
##  [901,] -0.40008209
##  [902,] -0.22391208
##  [903,]  1.89012803
##  [904,]  0.56885296
##  [905,] -0.84050711
##  [906,] -1.01667712
##  [907,]  0.74502297
##  [908,] -1.10476213
##  [909,]  0.65693796
##  [910,]  1.53778801
##  [911,]  1.53778801
##  [912,] -0.04774207
##  [913,] -0.84050711
##  [914,] -0.40008209
##  [915,] -1.10476213
##  [916,] -1.19284713
```

```
##  [917,] -1.36901714
##  [918,] -0.92859212
##  [919,] -1.01667712
##  [920,]  0.92119298
##  [921,] -0.75242211
##  [922,]  0.92119298
##  [923,] -1.10476213
##  [924,]  1.27353300
##  [925,] -0.84050711
##  [926,] -1.36901714
##  [927,] -0.40008209
##  [928,] -1.28093214
##  [929,]  0.30459795
##  [930,] -0.92859212
##  [931,] -0.48816710
##  [932,]  0.12842794
##  [933,]  0.92119298
##  [934,]  0.04034293
##  [935,]  0.65693796
##  [936,] -0.22391208
##  [937,] -0.57625210
##  [938,]  0.04034293
##  [939,] -1.10476213
##  [940,]  0.56885296
##  [941,]  0.74502297
##  [942,]  0.48076795
##  [943,]  0.48076795
##  [944,]  0.12842794
##  [945,]  1.89012803
##  [946,] -0.13582708
##  [947,] -0.92859212
##  [948,]  1.53778801
##  [949,] -0.31199709
##  [950,] -0.22391208
##  [951,] -1.01667712
##  [952,] -0.66433710
##  [953,]  0.30459795
##  [954,] -1.01667712
##  [955,] -0.84050711
##  [956,]  1.71395802
##  [957,]  2.24246805
##  [958,] -1.28093214
##  [959,] -0.84050711
##  [960,] -1.19284713
##  [961,]  2.85906308
##  [962,] -1.19284713
##  [963,]  1.97821303
##  [964,] -1.10476213
##  [965,]  0.12842794
##  [966,] -0.84050711
##  [967,] -0.75242211
##  [968,] -0.84050711
##  [969,] -1.10476213
##  [970,]  0.39268295
```

```
##  [971,] -0.22391208
##  [972,]  0.56885296
##  [973,]  0.12842794
##  [974,] -0.57625210
##  [975,] -1.10476213
##  [976,] -0.13582708
##  [977,] -0.31199709
##  [978,]  2.85906308
##  [979,] -0.92859212
##  [980,] -0.04774207
##  [981,] -1.19284713
##  [982,] -0.66433710
##  [983,] -0.48816710
##  [984,] -0.92859212
##  [985,] -0.13582708
##  [986,]  0.56885296
##  [987,]  0.83310797
##  [988,] -1.36901714
##  [989,] -1.10476213
##  [990,] -1.36901714
##  [991,] -0.75242211
##  [992,]  0.74502297
##  [993,]  1.36161800
##  [994,] -0.22391208
##  [995,] -0.66433710
##  [996,] -1.28093214
##  [997,]  0.83310797
##  [998,] -0.48816710
##  [999,]  1.44970300
## [1000,] -0.40008209
## attr(,"scaled:center")
## [1] 35.542
## attr(,"scaled:scale")
## [1] 11.35267
```

```
glimpse(df_credito_base)
```

```
## Rows: 1,000
## Columns: 21
## $ credit.rating                  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ account.balance                <dbl> 1, 1, 2, 1, 1, 1, 1, 1, 3, 2, 1, 1, ...
## $ credit.duration.months         <dbl> 18, 9, 12, 12, 12, 10, 8, 6, 18, 24,...
## $ previous.credit.payment.status <dbl> 3, 3, 2, 3, 3, 3, 3, 3, 3, 2, 3, 3, ...
## $ credit.purpose                 <dbl> 2, 4, 4, 4, 4, 4, 4, 4, 3, 3, 4, 1, ...
## $ credit.amount                  <dbl> 1049, 2799, 841, 2122, 2171, 2241, 3...
## $ savings                        <dbl> 1, 1, 2, 1, 1, 1, 1, 1, 1, 3, 1, 2, ...
## $ employment.duration            <dbl> 1, 2, 3, 2, 2, 1, 3, 1, 1, 1, 2, 3, ...
## $ installment.rate               <dbl> 4, 2, 2, 3, 4, 1, 1, 2, 4, 1, 2, 1, ...
## $ marital.status                 <dbl> 1, 3, 1, 3, 3, 3, 3, 3, 1, 1, 3, 4, ...
## $ guarantor                      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ residence.duration             <dbl> 4, 2, 4, 2, 4, 3, 4, 4, 4, 4, 2, 4, ...
## $ current.assets                 <dbl> 2, 1, 1, 1, 2, 1, 1, 1, 3, 4, 1, 3, ...
## $ age                            <dbl> 21, 36, 23, 39, 38, 48, 39, 40, 65, ...
## $ other.credits                  <dbl> 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, ...
## $ apartment.type                 <dbl> 1, 1, 1, 1, 2, 1, 2, 2, 2, 1, 1, 1, ...
## $ bank.credits                   <dbl> 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2, 2, ...
## $ occupation                     <dbl> 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 3, 3, ...
## $ dependents                     <dbl> 1, 2, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, ...
## $ telephone                      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ foreign.worker                 <dbl> 1, 1, 1, 2, 2, 2, 2, 2, 1, 1, 1, 1, ...
```

```
any(is.na(df_credito_base))
```

```
## [1] FALSE
```

```
# Pego as variáveis que serão normalizadas
vals_normalizar <-c('age', 'credit.amount','credit.duration.months')
head(vals_normalizar)
```

```
## [1] "age"                    "credit.amount"          "credit.duration.months"
```

```
# Pegos as variáveis que serão convertidas para factor.
vals_factor<- c('credit.rating','account.balance','previous.credit.payment.status','credit.purpo
se',
             'savings','employment.duration','installment.rate','marital.status','guarantor',
             'residence.duration','current.assets', 'other.credits','apartment.type','bank.credi
ts',
             'occupation','dependents','telephone','foreign.worker' )
```

```
# Crio as funções de normalização e transformação das variáveis
factorfunc <- function(df, var){
    for (i in var){
      df[[i]] = as.factor(df[[i]])
    }
  return(df)
}
normfunc <- function(df, var) {
  for (i in var){
    df[[i]] = scale(df[[i]],center = T, scale = T)
  }
  return(df)
}
```

```
# Faço a conversão das variaveis que deve ser fatores e normalizo as que precisam ser normalizad
as
# e jogo tudo no df_credito
df_credito <- factorfunc(df_credito_base, vals_factor)
df_credito <- normfunc(df_credito, vals_normalizar)
glimpse(df_credito)
```

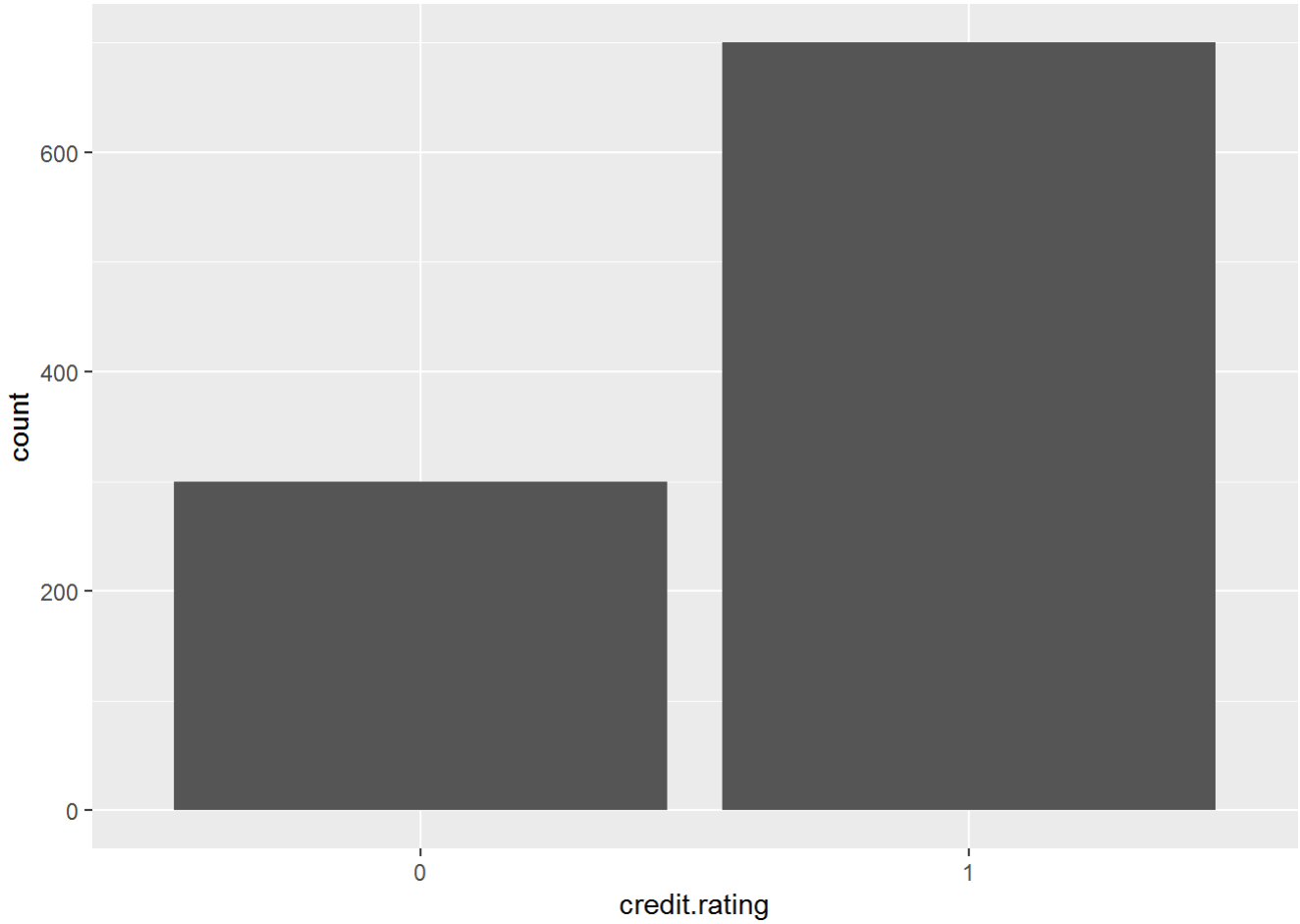
```
## Rows: 1,000
## Columns: 21
## $ credit.rating                  <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ account.balance                <fct> 1, 1, 2, 1, 1, 1, 1, 1, 3, 2, 1, 1, ...
## $ credit.duration.months         <dbl[,1]> <matrix[26 x 1]>
## $ previous.credit.payment.status <fct> 3, 3, 2, 3, 3, 3, 3, 3, 3, 2, 3, 3, ...
## $ credit.purpose                 <fct> 2, 4, 4, 4, 4, 4, 4, 4, 3, 3, 4, 1, ...
## $ credit.amount                  <dbl[,1]> <matrix[26 x 1]>
## $ savings                        <fct> 1, 1, 2, 1, 1, 1, 1, 1, 1, 3, 1, 2, ...
## $ employment.duration            <fct> 1, 2, 3, 2, 2, 1, 3, 1, 1, 1, 2, 3, ...
## $ installment.rate               <fct> 4, 2, 2, 3, 4, 1, 1, 2, 4, 1, 2, 1, ...
## $ marital.status                 <fct> 1, 3, 1, 3, 3, 3, 3, 3, 1, 1, 3, 4, ...
## $ guarantor                      <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ residence.duration             <fct> 4, 2, 4, 2, 4, 3, 4, 4, 4, 4, 2, 4, ...
## $ current.assets                 <fct> 2, 1, 1, 1, 2, 1, 1, 1, 3, 4, 1, 3, ...
## $ age                            <dbl[,1]> <matrix[26 x 1]>
## $ other.credits                  <fct> 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 2, 2, ...
## $ apartment.type                 <fct> 1, 1, 1, 1, 2, 1, 2, 2, 2, 1, 1, 1, ...
## $ bank.credits                   <fct> 1, 2, 1, 2, 2, 2, 2, 1, 2, 1, 2, 2, ...
## $ occupation                     <fct> 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 3, 3, ...
## $ dependents                     <fct> 1, 2, 1, 2, 1, 2, 1, 2, 1, 1, 2, 1, ...
## $ telephone                      <fct> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ foreign.worker                 <fct> 1, 1, 1, 2, 2, 2, 2, 2, 1, 1, 1, 1, ...
```

```
head(df_credito)
```

```
## # A tibble: 6 x 21
##   credit.rating account.balance credit.duration~ previous.credit~ credit.purpose
##   <fct>         <fct>                     <dbl> <fct>            <fct>
## 1 1             1                        -0.241 3                2
## 2 1             1                        -0.987 3                4
## 3 1             2                        -0.738 2                4
## 4 1             1                        -0.738 3                4
## 5 1             1                        -0.738 3                4
## 6 1             1                        -0.904 3                4
## # ... with 16 more variables: credit.amount[,1] <dbl>, savings <fct>,
## #   employment.duration <fct>, installment.rate <fct>, marital.status <fct>,
## #   guarantor <fct>, residence.duration <fct>, current.assets <fct>,
## #   age[,1] <dbl>, other.credits <fct>, apartment.type <fct>,
## #   bank.credits <fct>, occupation <fct>, dependents <fct>, telephone <fct>,
## #   foreign.worker <fct>
```

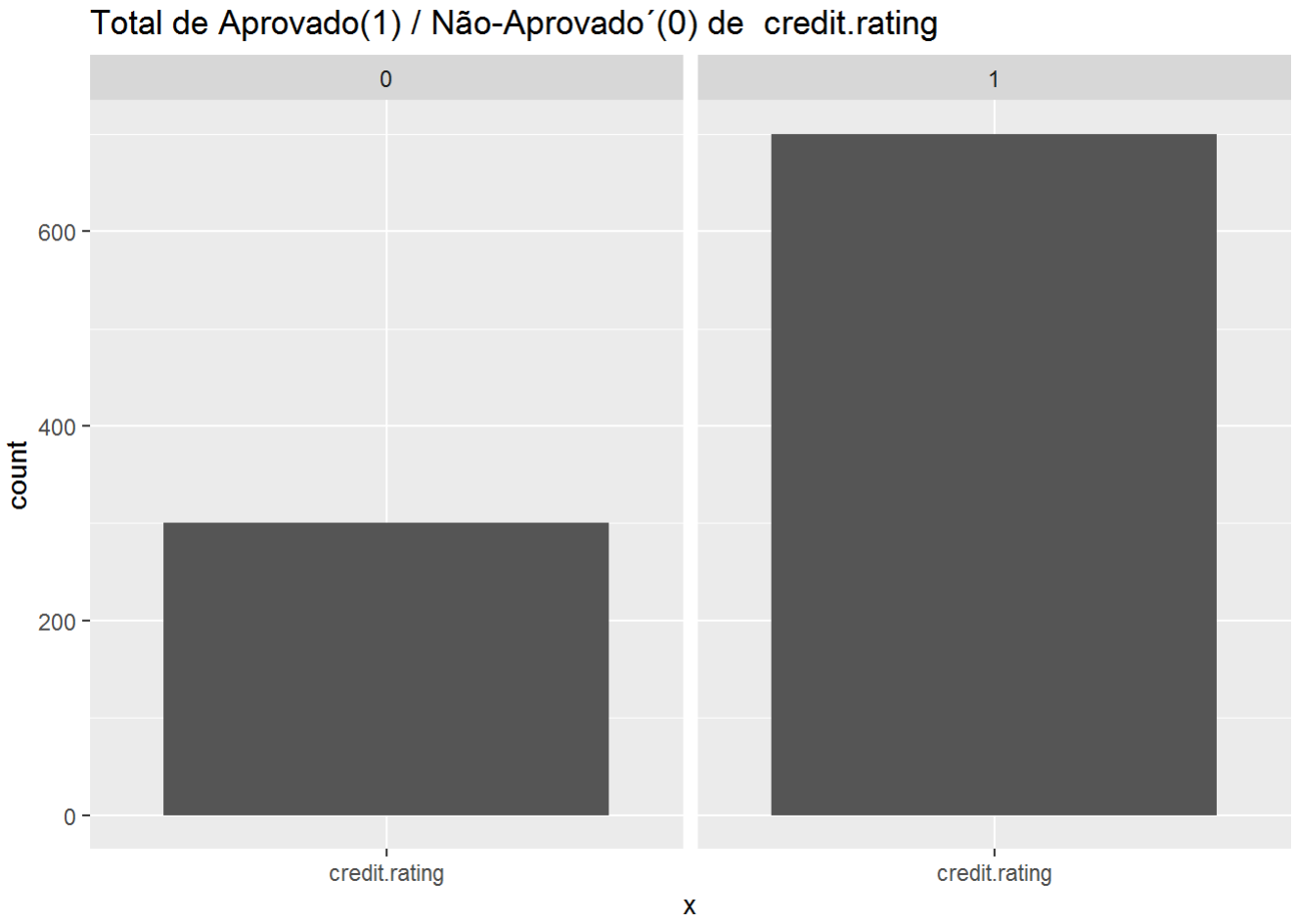
Analise Exploratória

```
# Verifico como esta o balanceamento da minha variável target, vejo que está desbalanceada
# temos muitos mais casos de aprovado do que negado, terei que balancear antes da criação dos mo
delos preditivos.
ggplot(df_credito, aes(x = credit.rating)) + geom_bar()
```

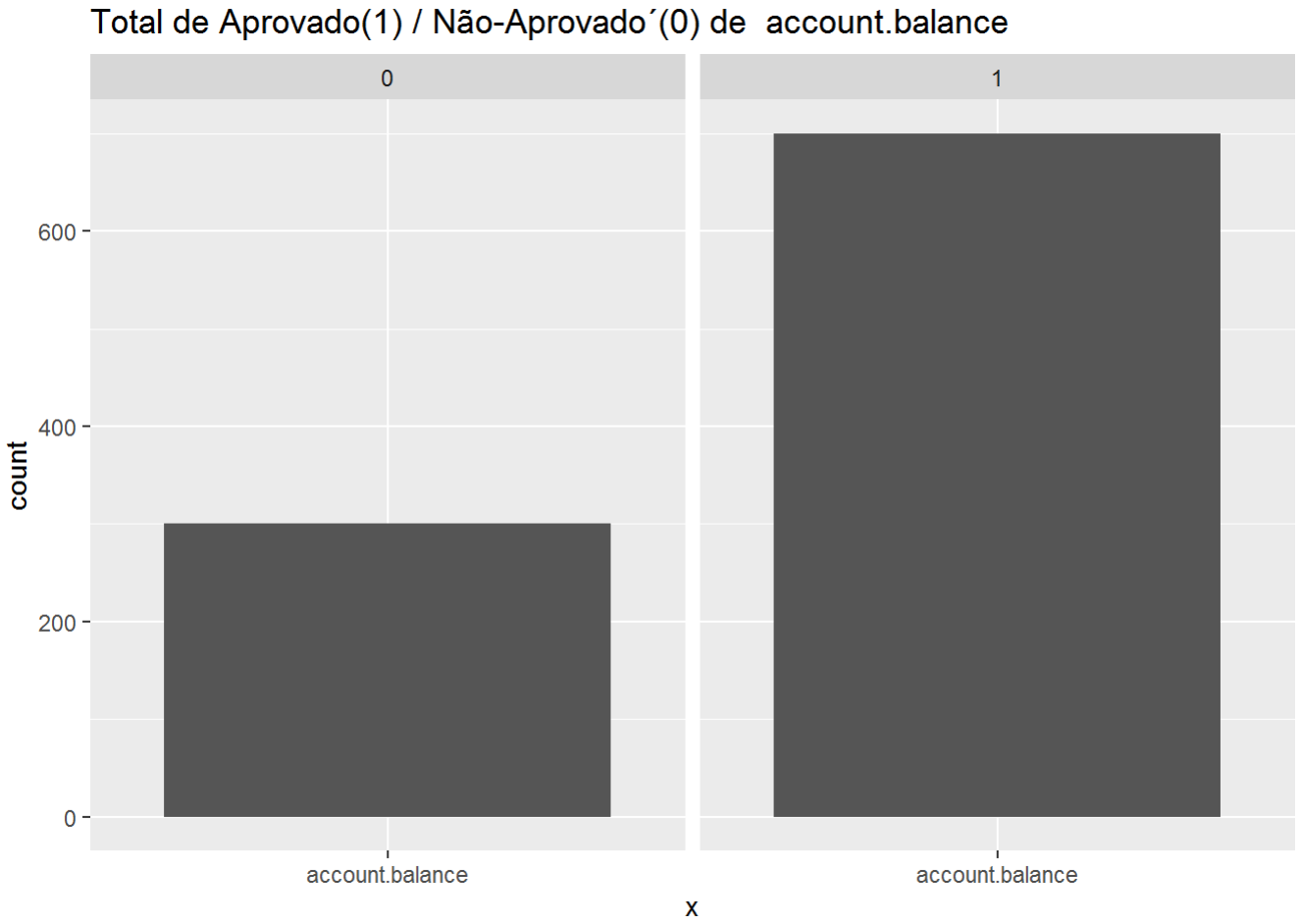


```
# Crio um plot para cada variável factor e mostro em um facet a quantidade de aprovados e reprov
ados
# por cada variavel.
# Primeiro crio a função depois passo ela para a lista de variavel.
plotfunction <- function(x){
    ggplot(df_credito, aes(x=x))+
    geom_bar()+
    facet_grid(. ~ credit.rating)+
    ggtitle(paste('Total de Aprovado(1) / Não-Aprovado´(0) de ',x))
}

lapply(vals_factor, plotfunction)
```

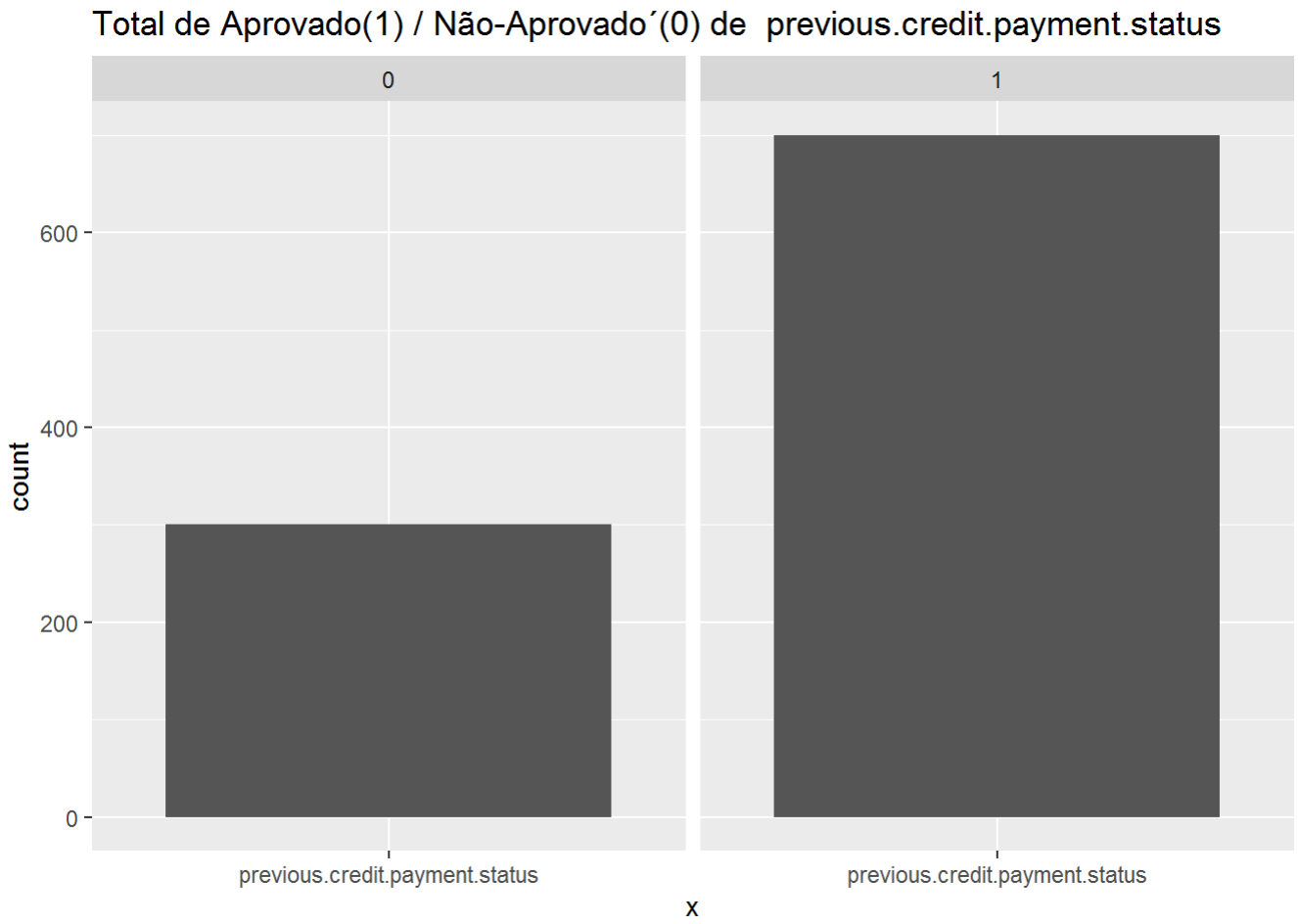
```
## [[1]]
```



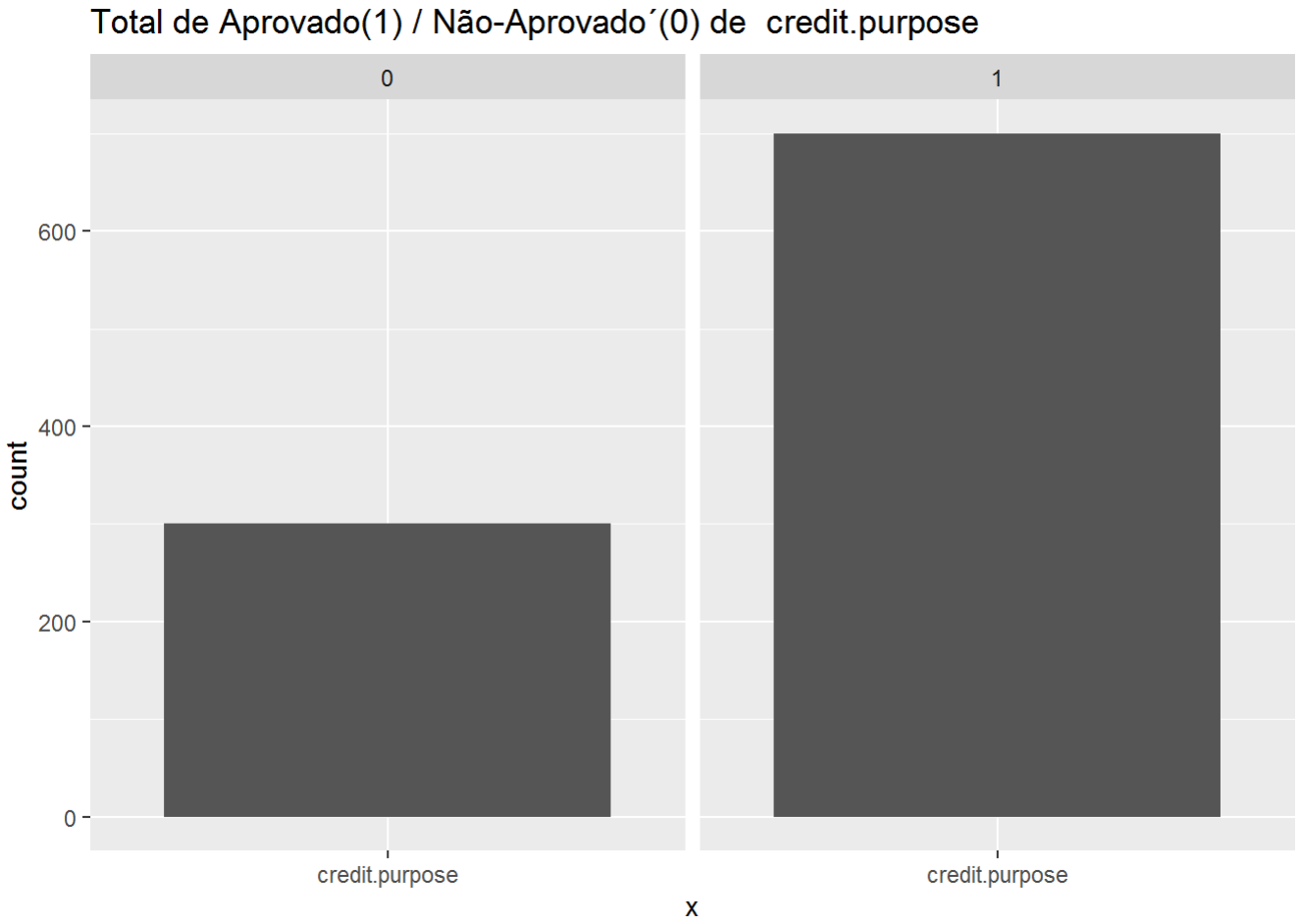

```
##
## [[2]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de account.balance
0
1
600
400
200
0
count
account.balance
account.balance
x

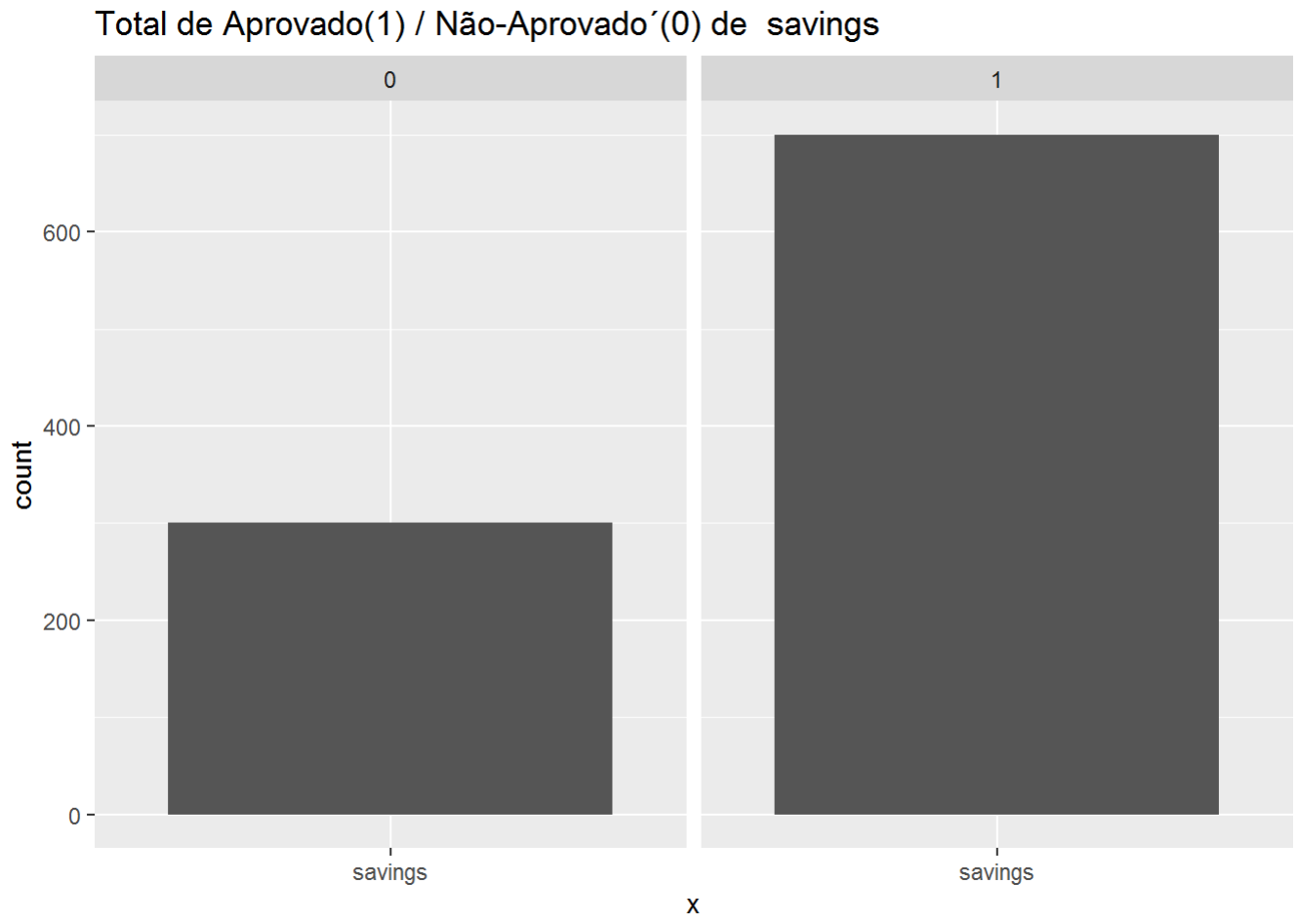
```
##
## [[3]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de previous.credit.payment.status
0
1
600
400
200
0
count
previous.credit.payment.status
previous.credit.payment.status
x

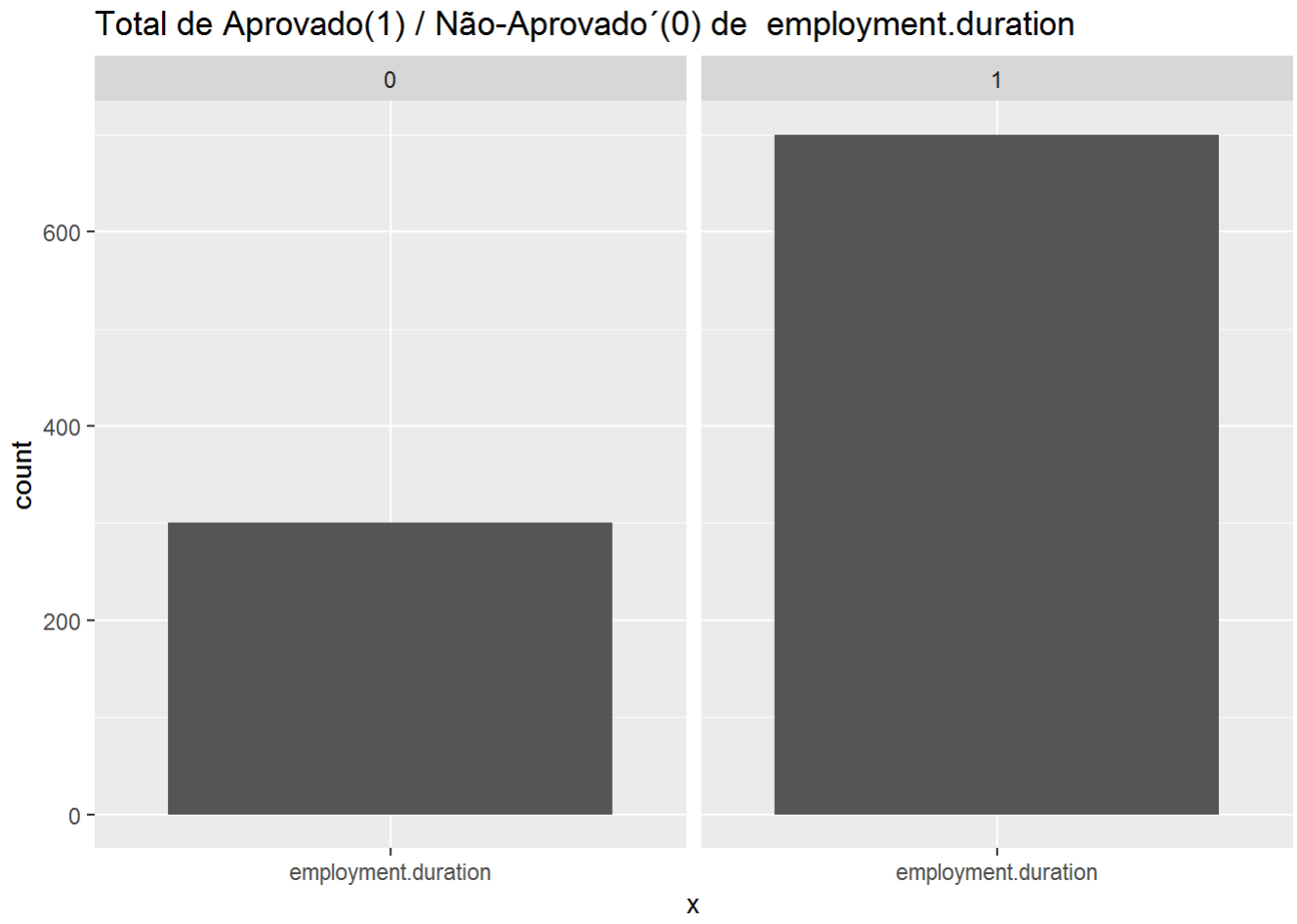
```
##
## [[4]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de credit.purpose
0
1
600
400
200
0
count
credit.purpose
credit.purpose
x

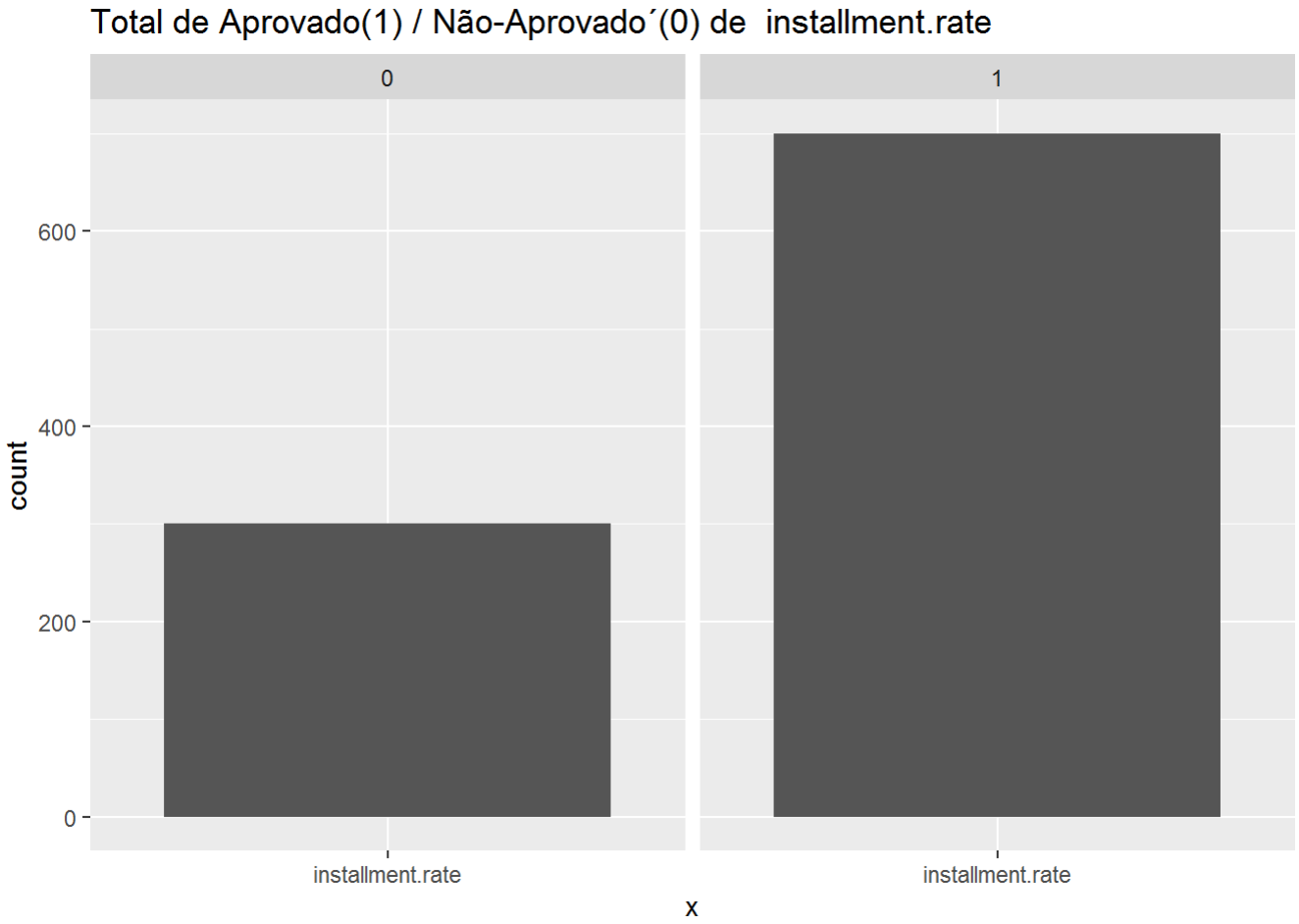
```
##
## [[5]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de savings
0
1
600
400
200
0
count
savings
savings
x

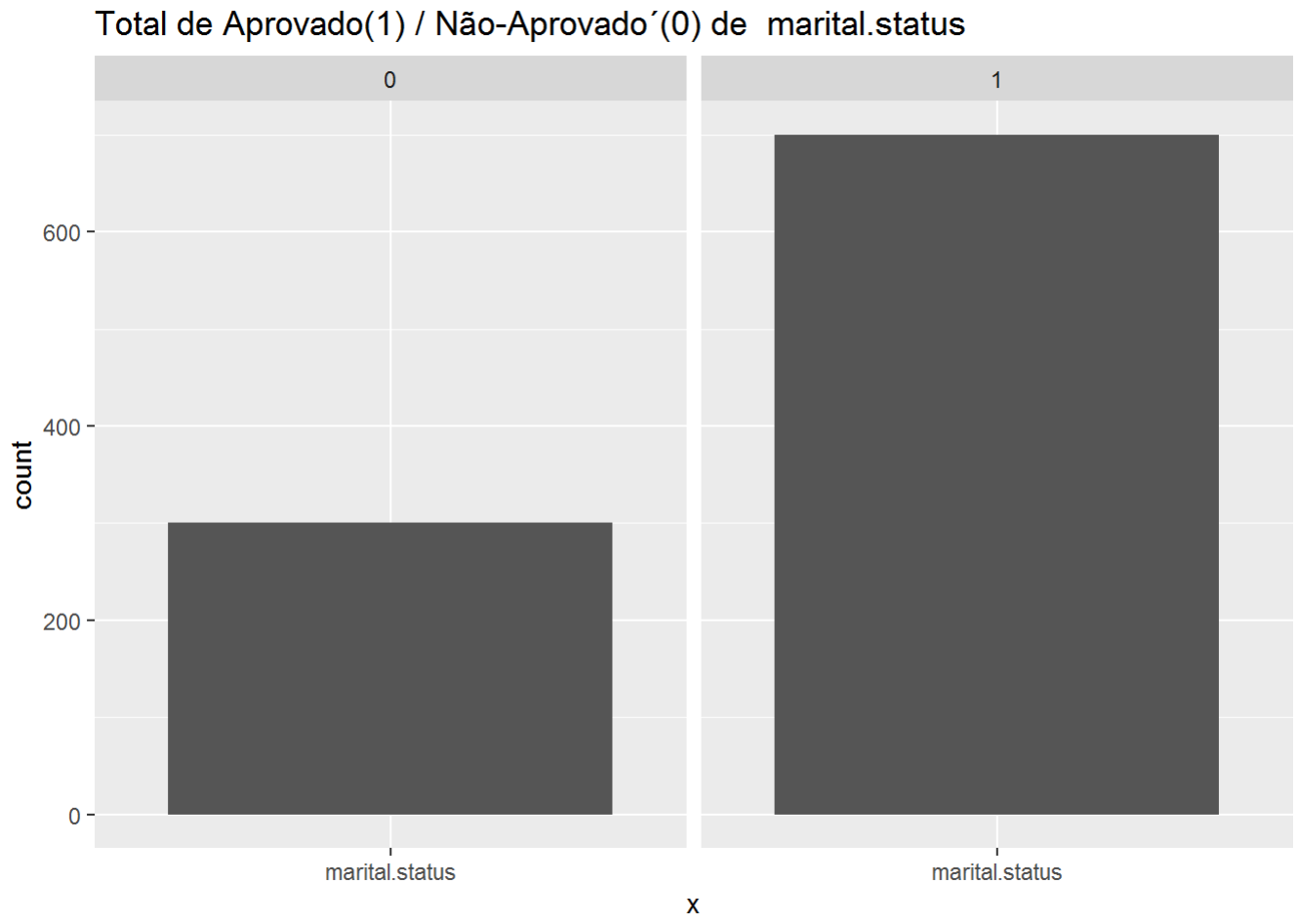
```
##
## [[6]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de employment.duration
0
1
600
400
200
0
count
employment.duration
employment.duration
x

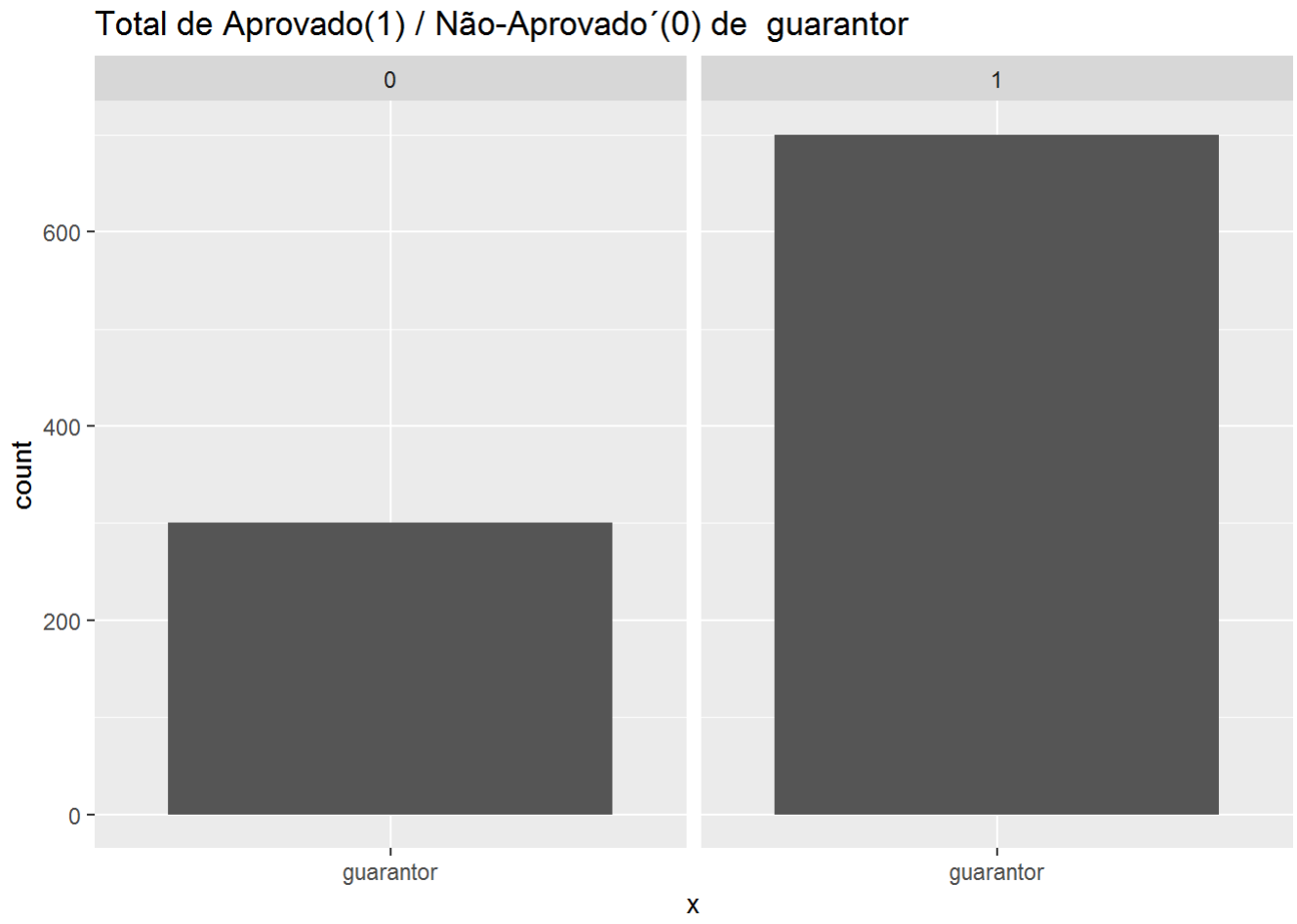
```
##
## [[7]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de installment.rate
0
1
600
400
200
0
count
installment.rate
installment.rate
x

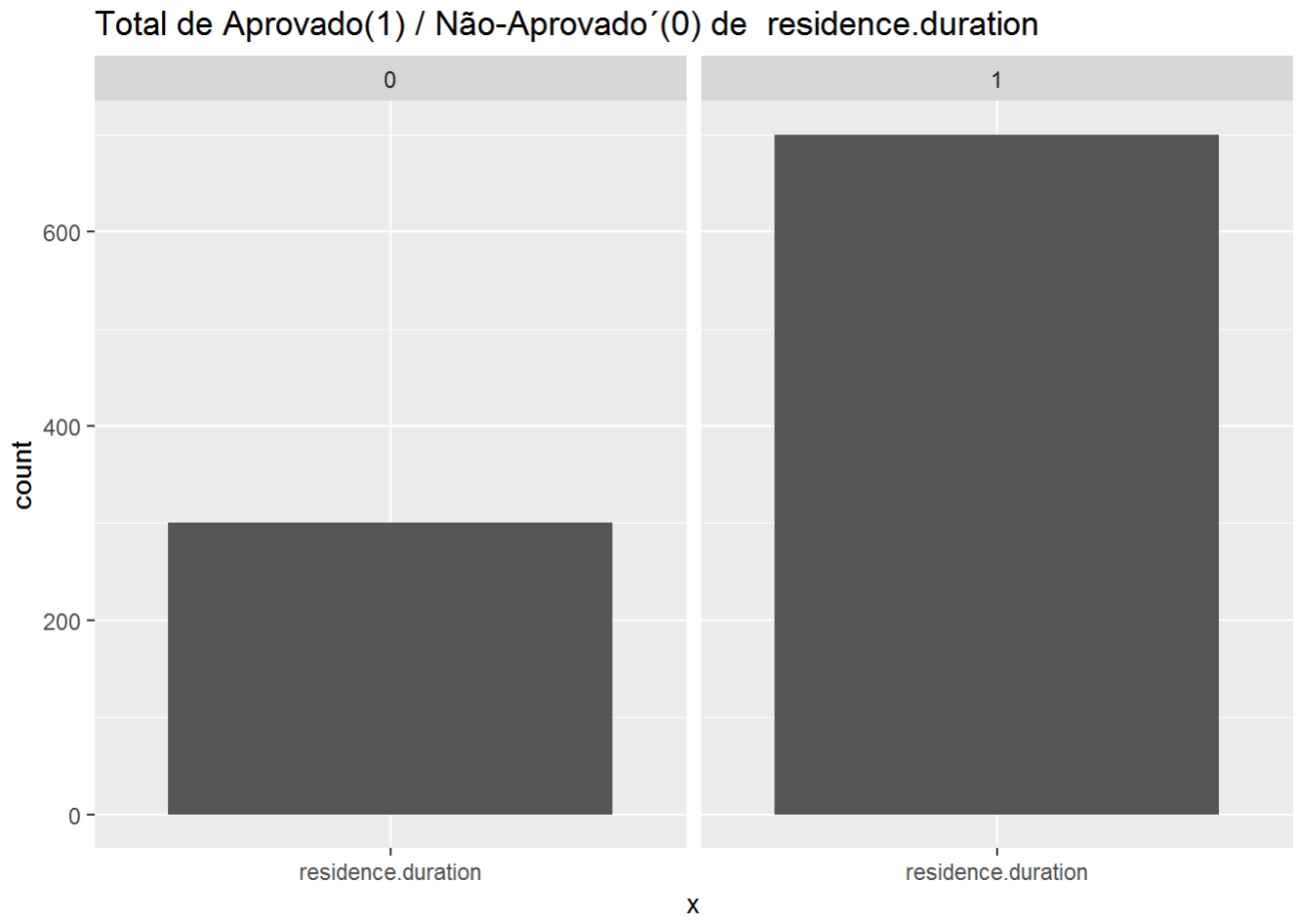
```
##
## [[8]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de marital.status
0
1
600
400
200
0
count
marital.status
marital.status
x

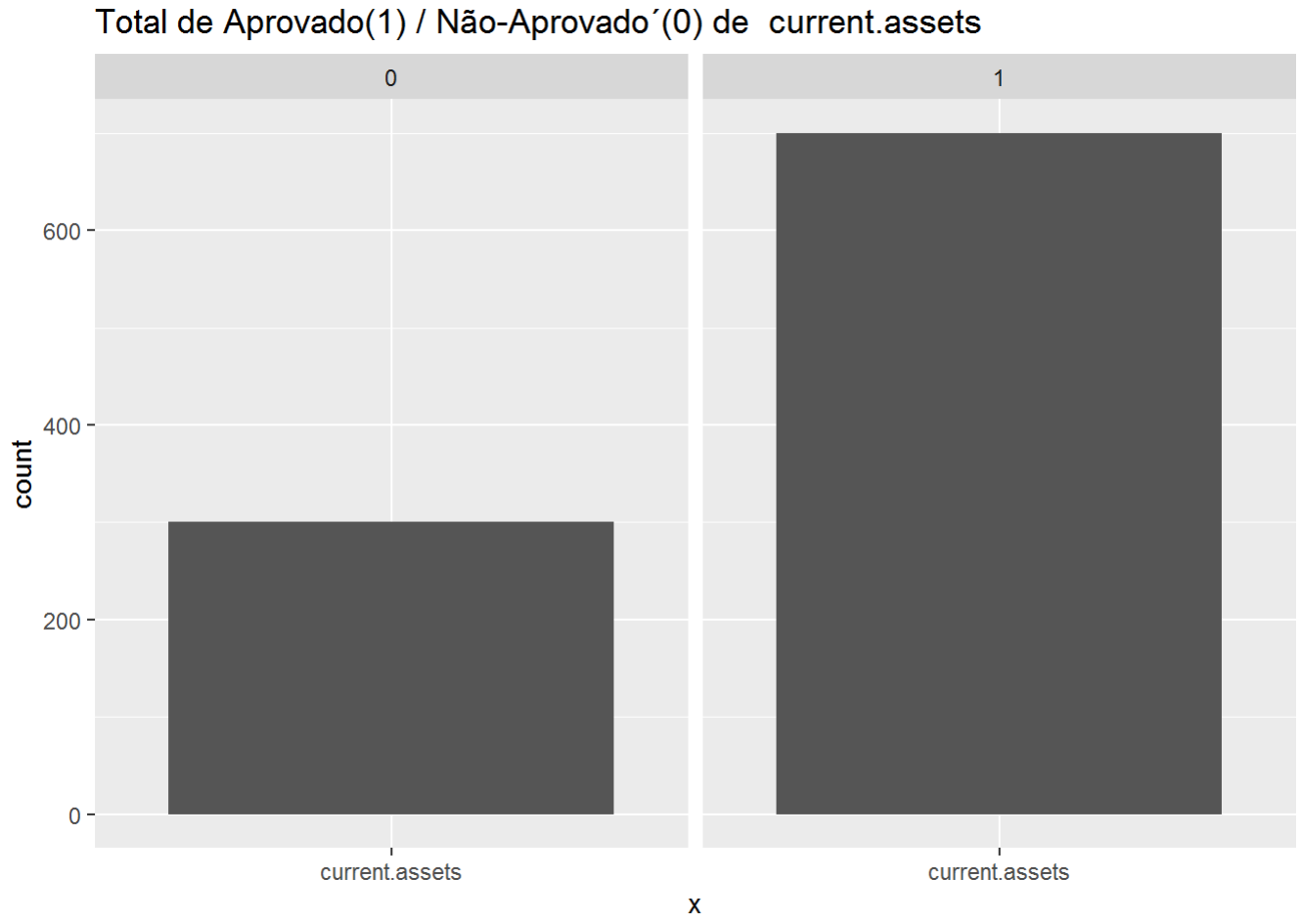
```
##
## [[9]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de guarantor
0
1
600
400
200
0
count
guarantor
guarantor
x

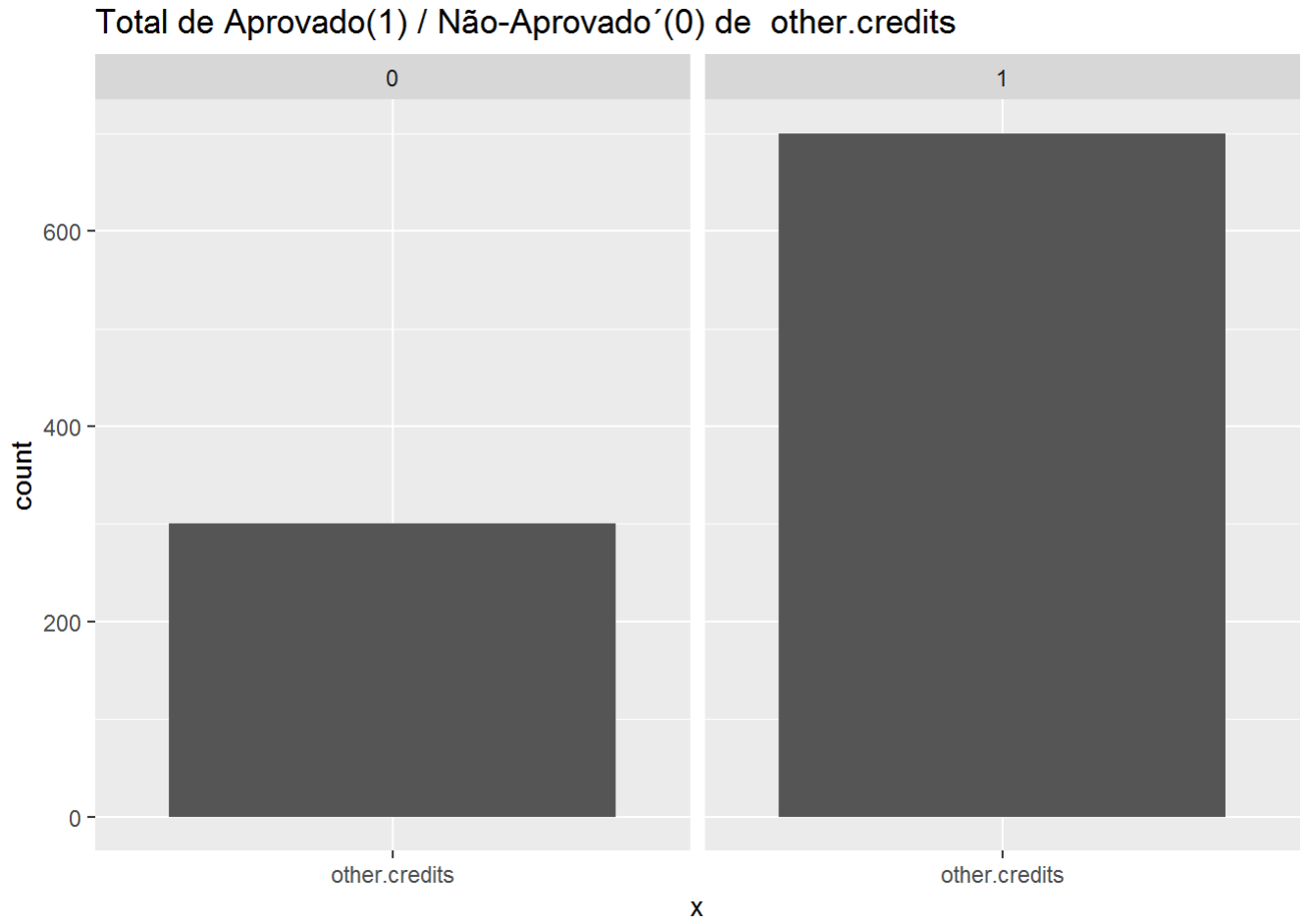
```
##
## [[10]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de residence.duration
0
1
600
400
200
0
count
residence.duration
residence.duration
x

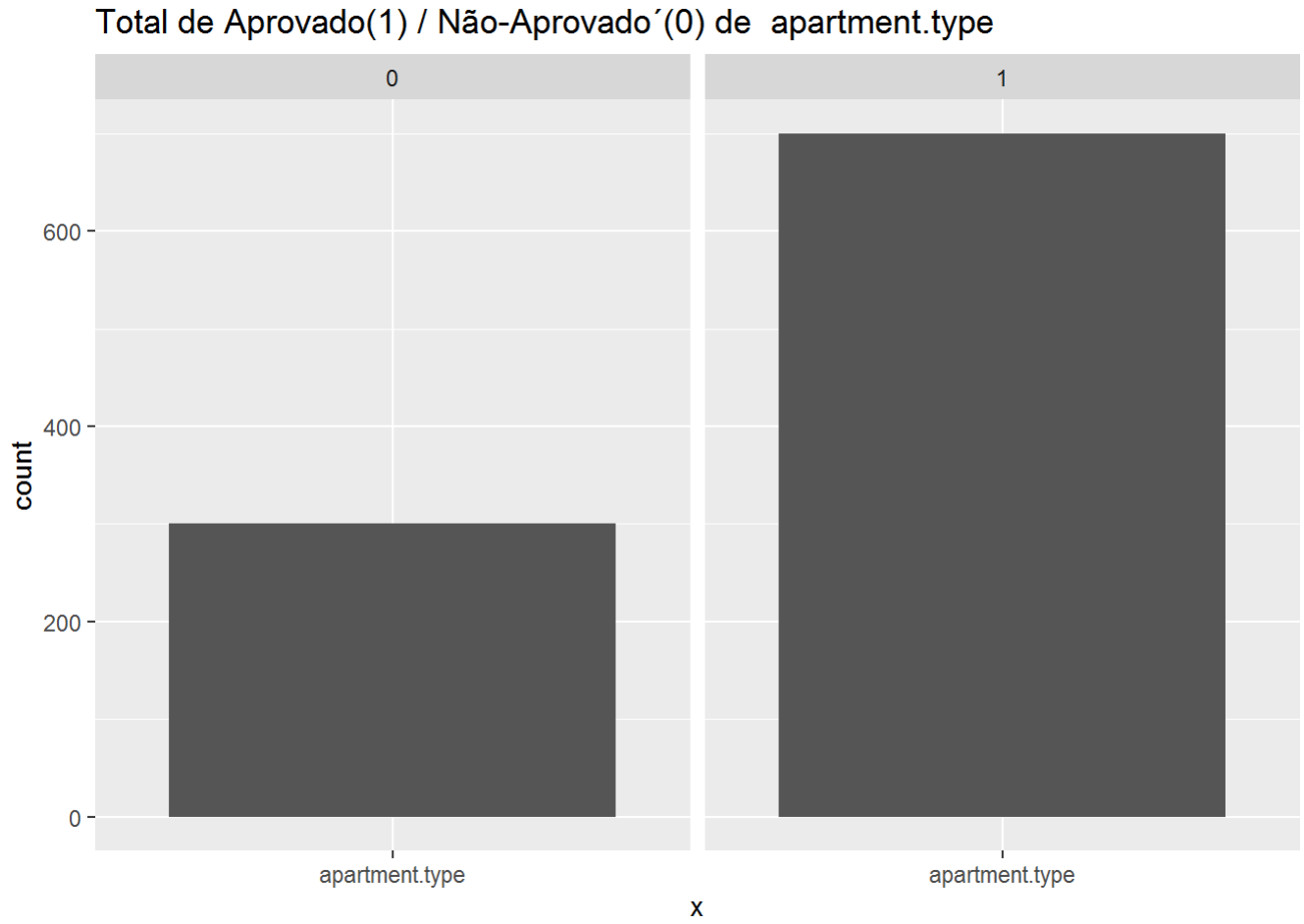
```
##
## [[11]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de current.assets
0
1
600
400
200
0
count
current.assets
current.assets
x

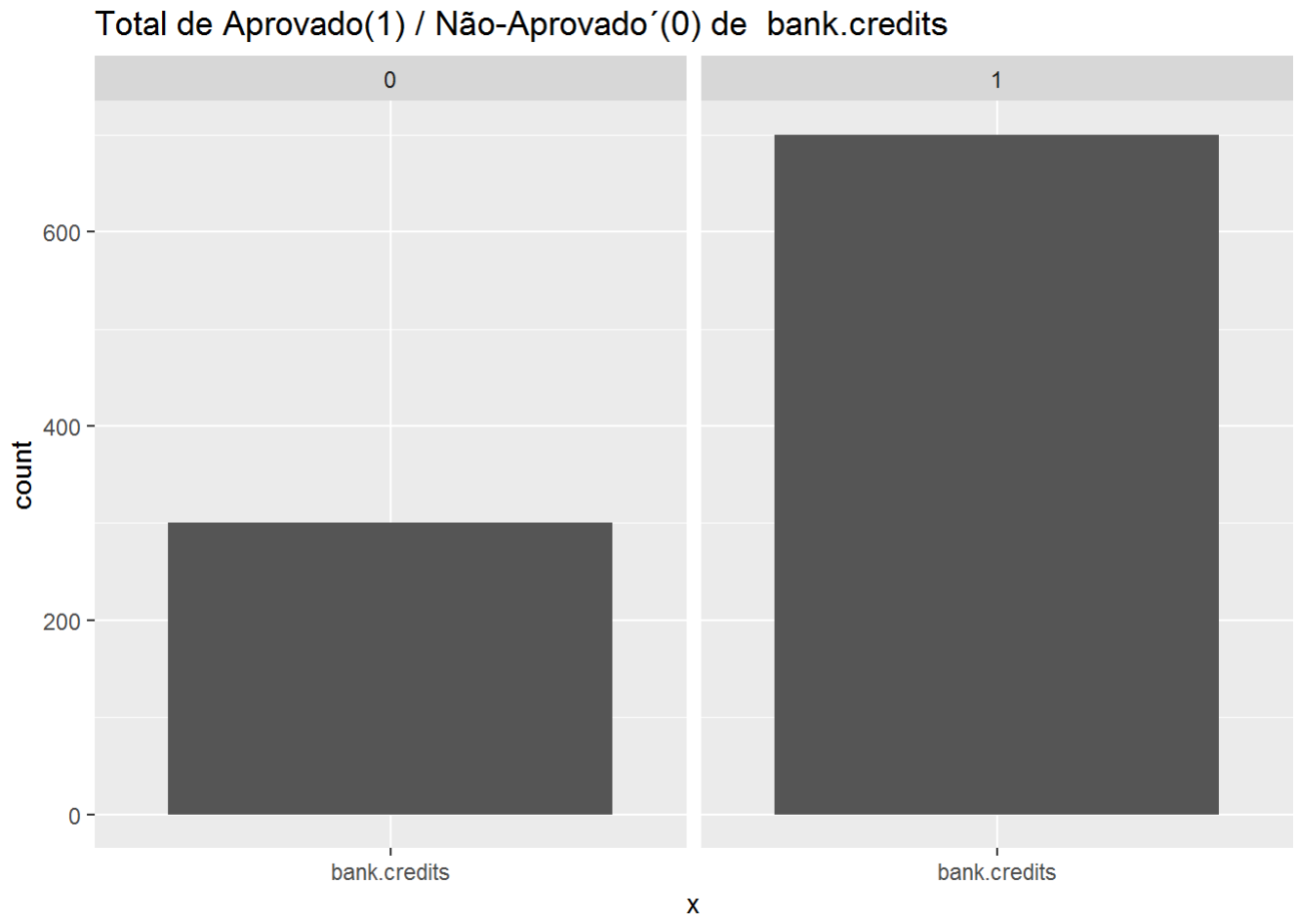
```
##
## [[12]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de other.credits
0
1
600
400
200
0
count
other.credits
other.credits
x

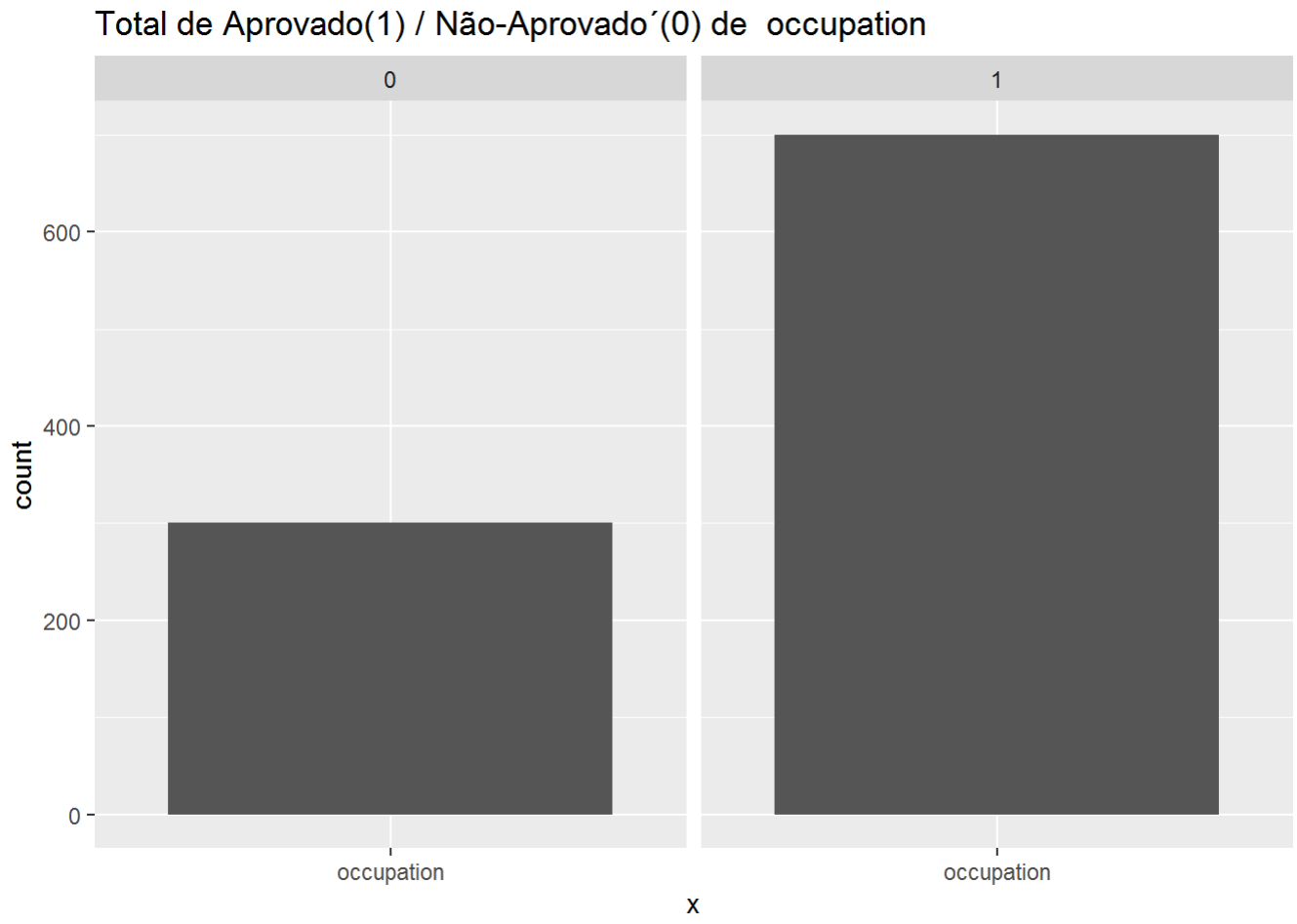
```
##
## [[13]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de apartment.type
0
1
600
400
200
0
count
apartment.type
apartment.type
x

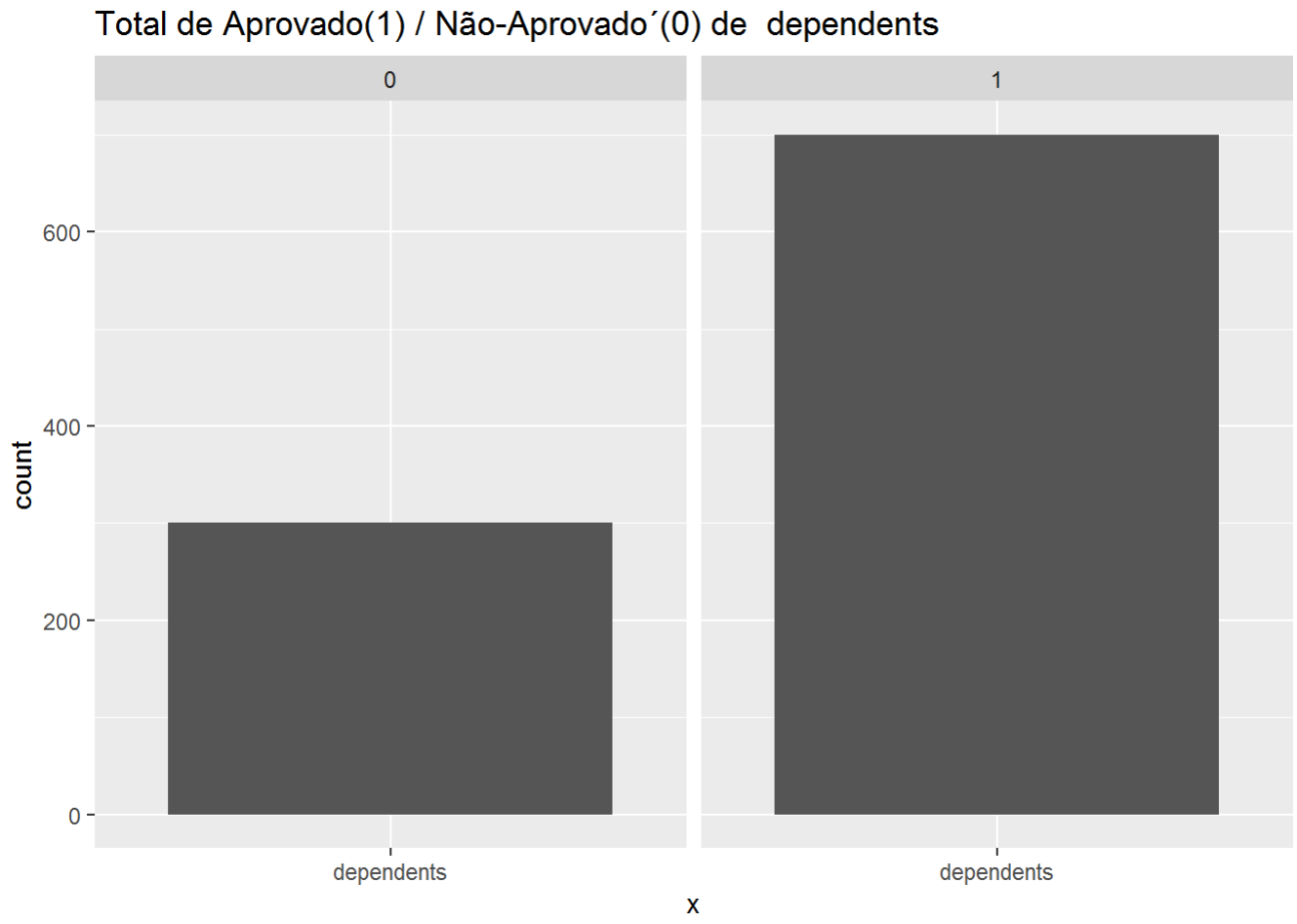
```
##
## [[14]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de bank.credits
0
1
600
400
200
0
count
bank.credits
bank.credits
x

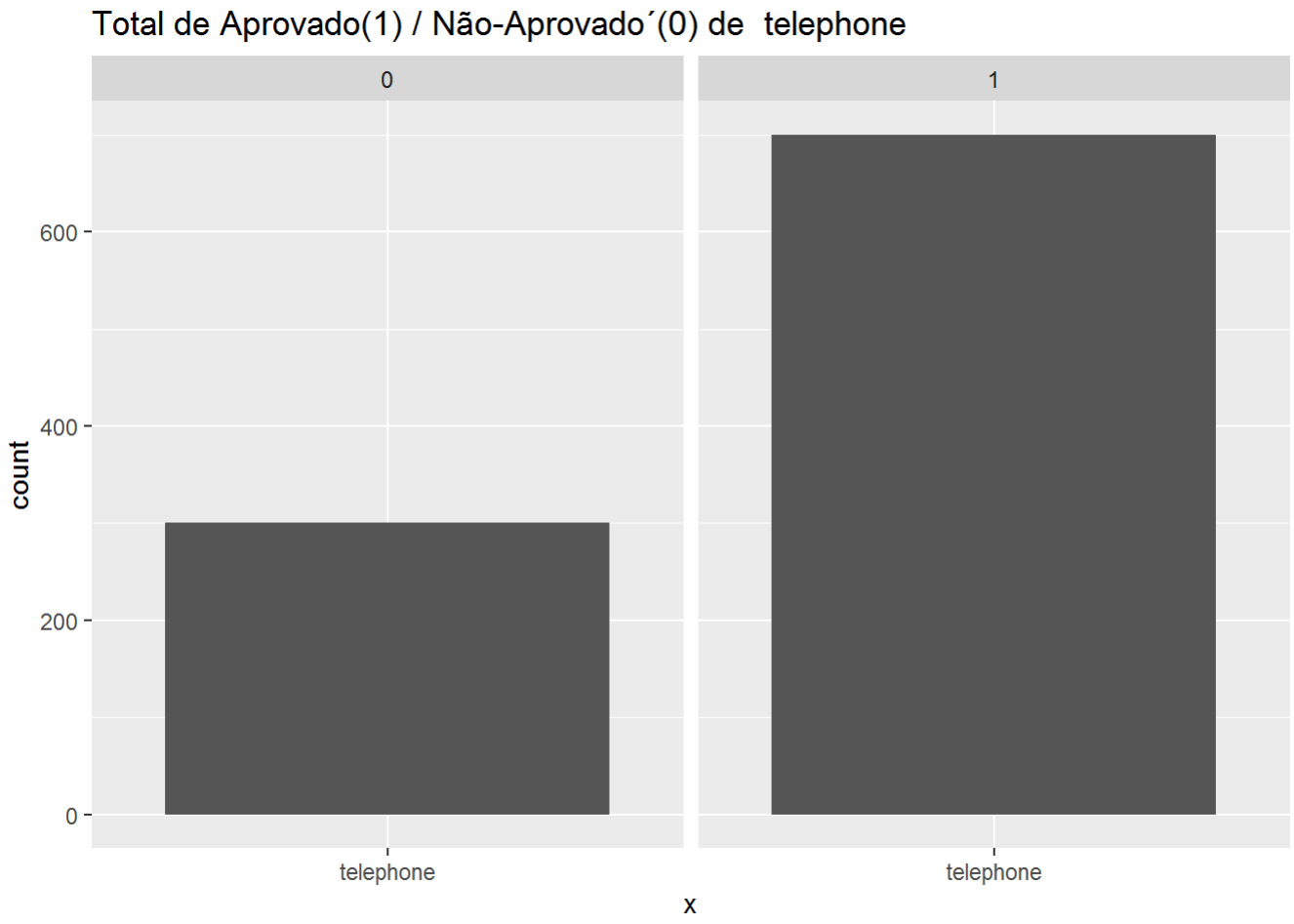
```
##
## [[15]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de occupation
0
1
600
400
200
0
count
occupation
occupation
x

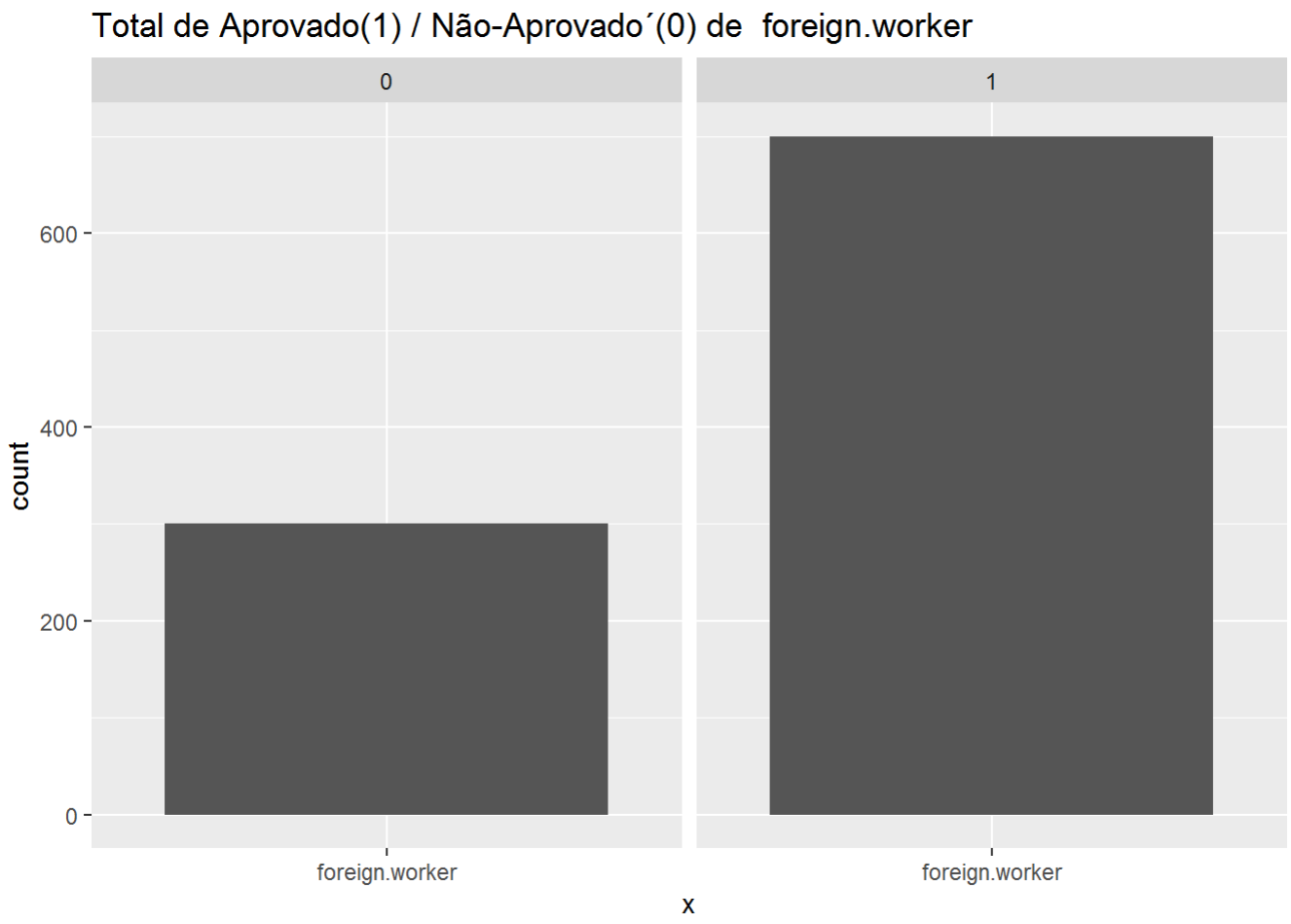
```
##
## [[16]]
```


Total de Aprovado(1) / Não-Aprovado´(0) de dependents
0
1
600
400
200
0
count
dependents
dependents
x

```
##
## [[17]]
```

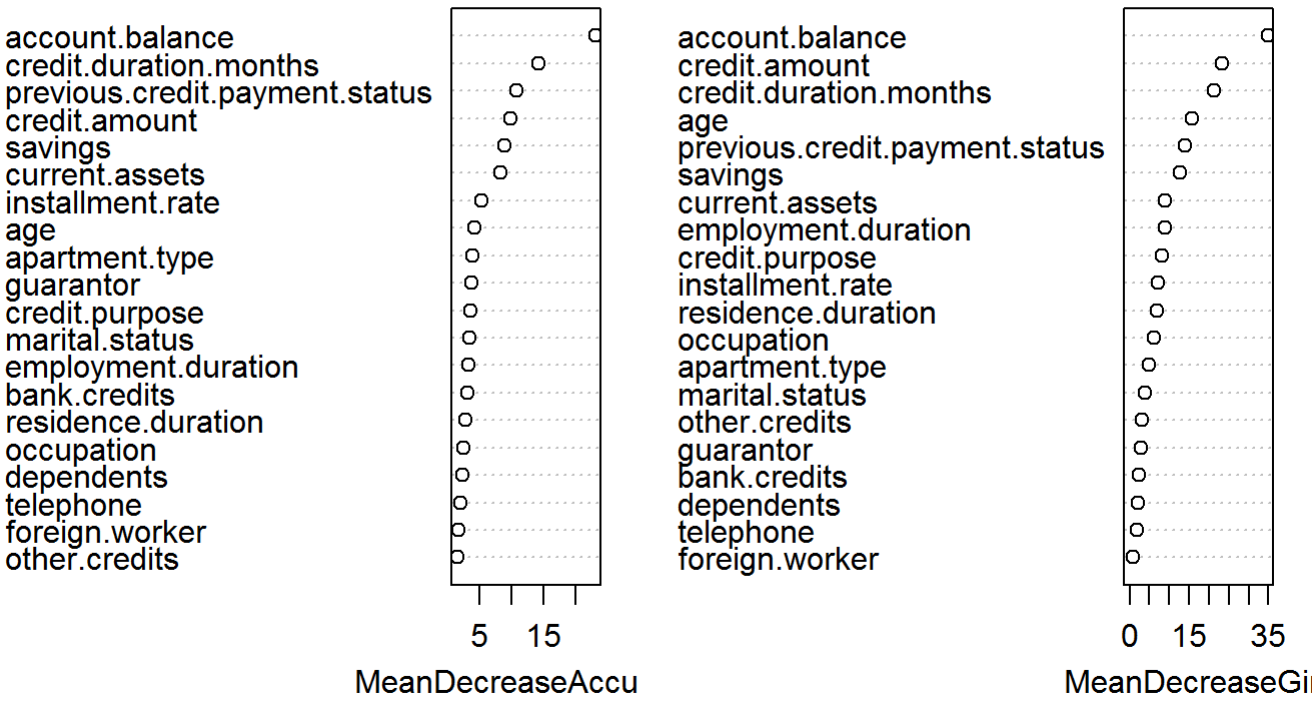

Total de Aprovado(1) / Não-Aprovado´(0) de telephone
0
1
600
400
200
0
count
telephone
telephone
x

```
##
## [[18]]
```



Feature Selection (Seleção de Variáveis)

```
# Uso o randomForest para a seleção das variaveis mais importante para o modelo

modimportance <- randomForest(data=df_credito, credit.rating ~.,
                              ntree = 200, nodesize = 20, importance = T)

varImpPlot(modimportance)
```



Split dos dados

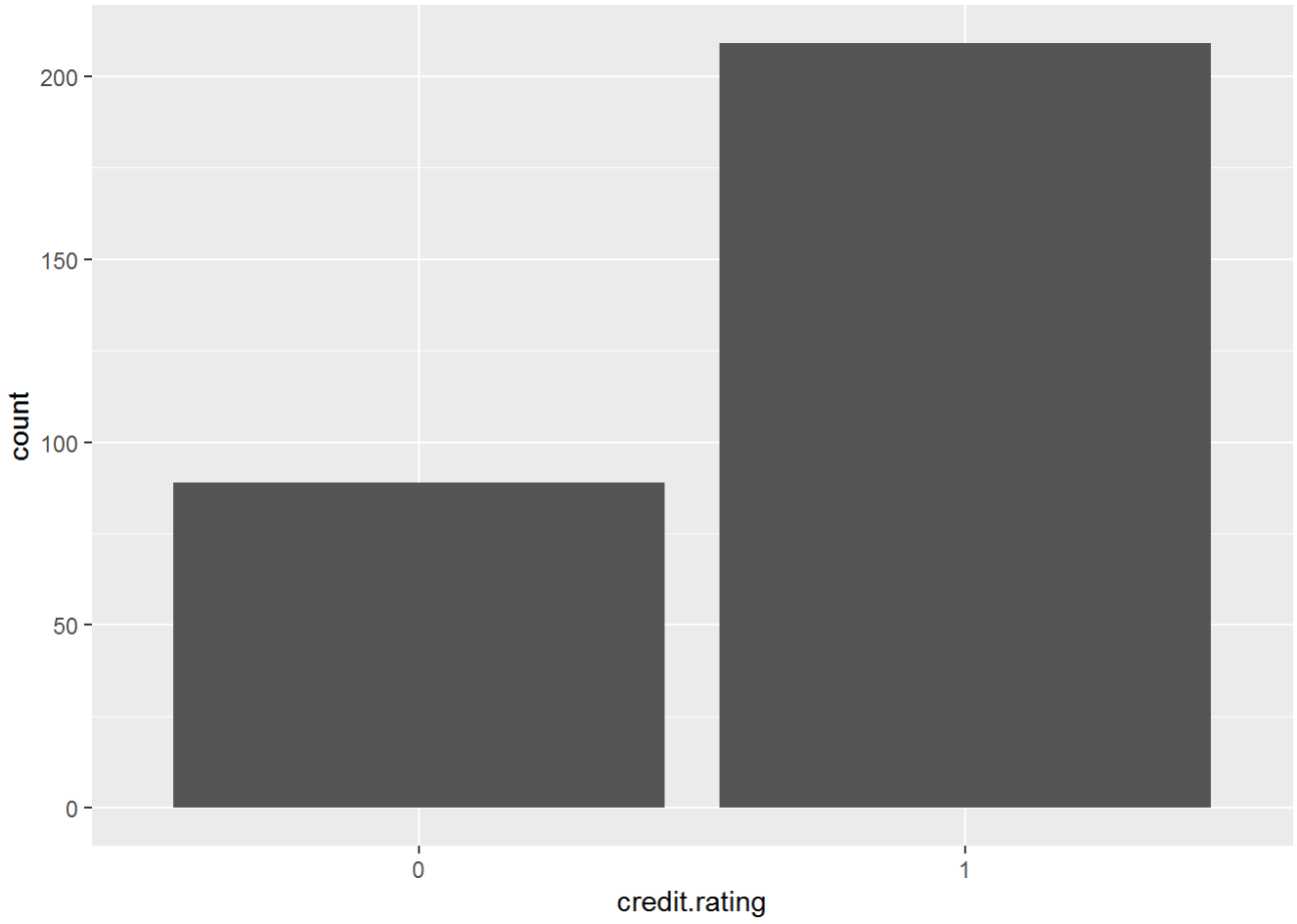
```
# Crio os dados de treino e teste para o treinamento dos algoritmos

split = sample.split(df_credito$age, SplitRatio = 0.70)

treino = subset(df_credito, split == TRUE)
teste = subset(df_credito, split == FALSE)

# Verifico as dimensões dos dados de treino e teste.
dim(treino)
```

```
## [1] 702  21
```

```
dim(teste)
```

```
## [1] 298  21
```

```
ggplot(teste, aes(x = credit.rating)) + geom_bar()
```

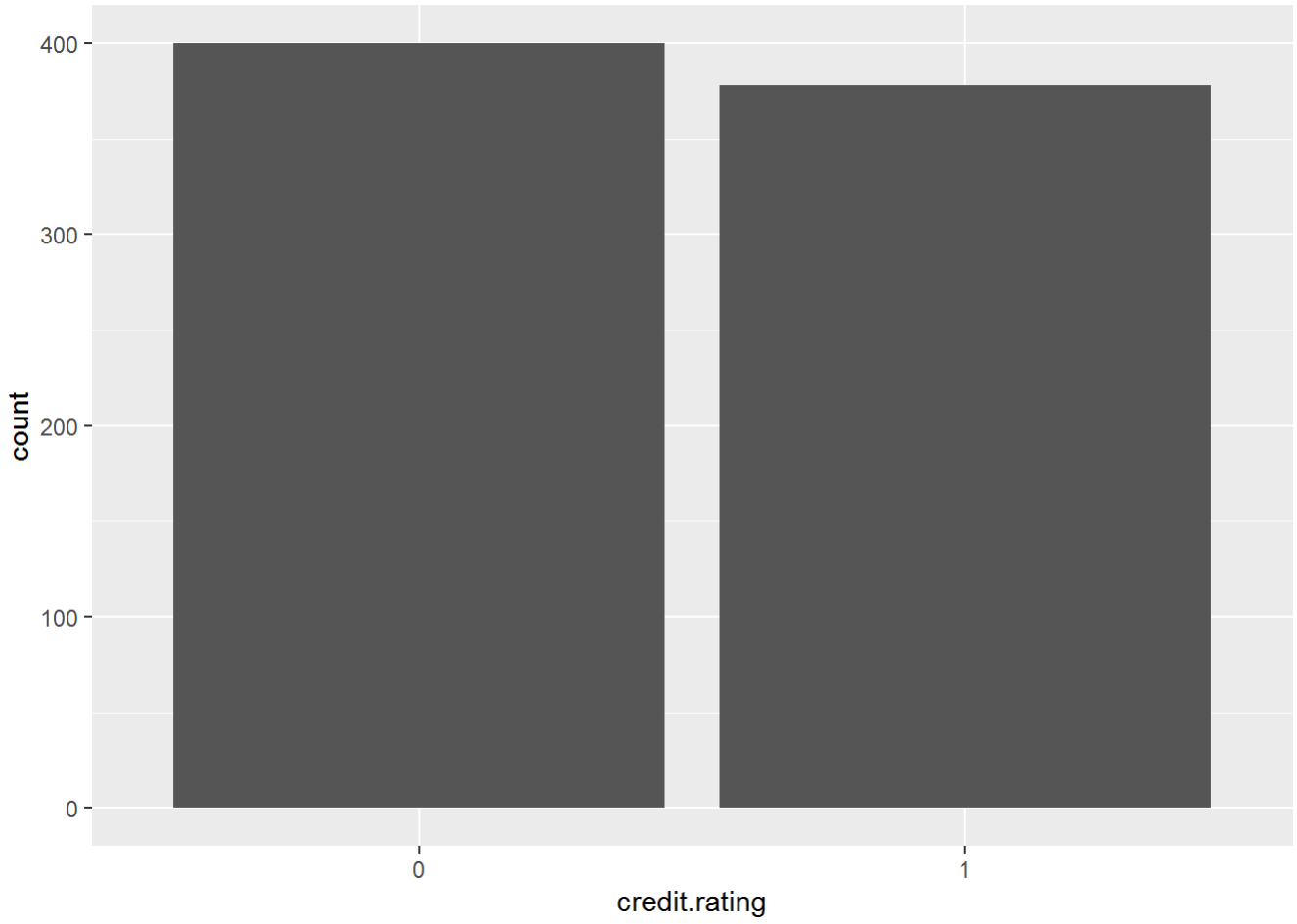


Balanceamento dos dados

```
# Para usar o SMOTE temos que converter o dataset treino pra data.frame, pois ao carregar com o
 readr
# ele ja tras como tibble e o smote da erro. Fiz a correção balancei e como mostrado no plot,
# temos agora uma variavel target balanceada, podemos então treinar os modelos.

treino_balanced <- as.data.frame(treino)
class(treino_balanced)
```

```
## [1] "data.frame"
```

```
treino_balanced <- SMOTE(credit.rating ~ .,treino_balanced, perc.over = 90, perc.under = 200)
ggplot(treino_balanced, aes(x = credit.rating)) + geom_bar()
```



Algoritmos de aprendizagem

```
# Modelo 1 com o KSVM  library(kernlab)

library(kernlab)
```

```
##
## Attaching package: 'kernlab'
```

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

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

```
modelo_v1 <- ksvm(credit.rating ~ account.balance
                  +credit.duration.months
                  +previous.credit.payment.status
                  +savings
                  +current.assets
                  ,data= treino_balanced,type="C-bsvc", kernel = "rbfdot")

previsao_v1 <- predict(modelo_v1, teste)
confusionMatrix(previsao_v1,teste$credit.rating)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0  69  77
##          1  20 132
##
##                Accuracy : 0.6745
##                  95% CI : (0.6181, 0.7274)
##     No Information Rate : 0.7013
##     P-Value [Acc > NIR] : 0.8588
##
##                   Kappa : 0.3437
##
##  Mcnemar's Test P-Value : 1.301e-08
##
##             Sensitivity : 0.7753
##             Specificity : 0.6316
##          Pos Pred Value : 0.4726
##          Neg Pred Value : 0.8684
##              Prevalence : 0.2987
##          Detection Rate : 0.2315
##    Detection Prevalence : 0.4899
##       Balanced Accuracy : 0.7034
##
##        'Positive' Class : 0
##
```

```
# Modelo 2 com o RandomForest library('randomForest')
library('randomForest')
?randomForest
```

```
## starting httpd help server ... done
```

```
modelo_v2 <- randomForest(credit.rating ~ account.balance
                          +credit.duration.months
                          +previous.credit.payment.status
                          +savings
                          +current.assets
                          ,data= treino_balanced
                          ,ntree = 100
                          ,nodesize = 10
                          ,cost = Cost_func)

previsao_v2 <- predict(modelo_v2, teste)
confusionMatrix(previsao_v2, teste$credit.rating)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0  65  77
##          1  24 132
##
##                Accuracy : 0.6611
##                  95% CI : (0.6043, 0.7147)
##     No Information Rate : 0.7013
##     P-Value [Acc > NIR] : 0.9419
##
##                   Kappa : 0.3091
##
##  Mcnemar's Test P-Value : 2.289e-07
##
##             Sensitivity : 0.7303
##             Specificity : 0.6316
##          Pos Pred Value : 0.4577
##          Neg Pred Value : 0.8462
##              Prevalence : 0.2987
##          Detection Rate : 0.2181
##    Detection Prevalence : 0.4765
##       Balanced Accuracy : 0.6810
##
##        'Positive' Class : 0
##
```

```r
# Modelo3 svm do pacote library(e1071)
library(e1071)
modelo_v3 <- svm(credit.rating ~ account.balance
                 +credit.duration.months
                 +previous.credit.payment.status
                 +savings
                 +current.assets
                 ,data= treino_balanced
                 ,type = 'C-classification')

previsao_v3 <- predict(modelo_v3, teste)
confusionMatrix(previsao_v3, teste$credit.rating)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0  71  78
##          1  18 131
##
##                Accuracy : 0.6779
##                  95% CI : (0.6215, 0.7306)
##     No Information Rate : 0.7013
##     P-Value [Acc > NIR] : 0.829
##
##                   Kappa : 0.3557
##
##  Mcnemar's Test P-Value : 1.726e-09
##
##             Sensitivity : 0.7978
##             Specificity : 0.6268
##          Pos Pred Value : 0.4765
##          Neg Pred Value : 0.8792
##              Prevalence : 0.2987
##          Detection Rate : 0.2383
##    Detection Prevalence : 0.5000
##       Balanced Accuracy : 0.7123
##
##        'Positive' Class : 0
##
```

```r
#Modelo 4  naiveBayes do  pacote library(e1071)
library(e1071)
?naiveBayes
modelo_v4 <- naiveBayes(credit.rating ~ account.balance
                  +credit.duration.months
                  +previous.credit.payment.status
                  +savings
                  +current.assets
                  ,data= treino_balanced
                   )
previsao_v4 <- predict(modelo_v4, teste)
confusionMatrix(previsao_v4,teste$credit.rating)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0  62  54
##          1  27 155
##
##                Accuracy : 0.7282
##                  95% CI : (0.6739, 0.7779)
##     No Information Rate : 0.7013
##     P-Value [Acc > NIR] : 0.171436
##
##                   Kappa : 0.4031
##
##  Mcnemar's Test P-Value : 0.003866
##
##             Sensitivity : 0.6966
##             Specificity : 0.7416
##          Pos Pred Value : 0.5345
##          Neg Pred Value : 0.8516
##              Prevalence : 0.2987
##          Detection Rate : 0.2081
##    Detection Prevalence : 0.3893
##       Balanced Accuracy : 0.7191
##
##        'Positive' Class : 0
##
```

```r
# Tentando Otimizar o modelo com o C5.0 e library(C50)
library(C50)
Cost_func <- matrix(c(0, 2, 1.5, 0), nrow = 2, dimnames = list(c("0", "1"), c("0", "1")))

modelo_v9 <- C5.0(credit.rating ~ account.balance
                  +credit.duration.months
                  +previous.credit.payment.status
                  +savings
                  ,data= treino_balanced
                  ,trials = 100,
                  cost = Cost_func)


modelo_v9 <- predict(modelo_v9, teste)
confusionMatrix(modelo_v9,teste$credit.rating)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##          0  72  73
##          1  17 136
##
##                Accuracy : 0.698
##                  95% CI : (0.6424, 0.7496)
##     No Information Rate : 0.7013
##     P-Value [Acc > NIR] : 0.5785
##
##                   Kappa : 0.3894
##
##  Mcnemar's Test P-Value : 6.731e-09
##
##             Sensitivity : 0.8090
##             Specificity : 0.6507
##          Pos Pred Value : 0.4966
##          Neg Pred Value : 0.8889
##              Prevalence : 0.2987
##          Detection Rate : 0.2416
##    Detection Prevalence : 0.4866
##       Balanced Accuracy : 0.7299
##
##        'Positive' Class : 0
##
```