

atividade 3

Arthur Vinicius

2023-10-05

Roteiro do trabalho:

- 1) Escolher um banco de dados;
- 2) Classificar as variáveis deste banco de dados;
- 3) Escolher uma das variáveis quantitativas e construir a tabela de frequência pontual ou intervalar.
- 4) Calcular as medidas-resumo.
- 5) Construir os gráficos: histograma e boxplot.

Banco de dados utilizado: <https://github.com/fivethirtyeight/data/tree/master/alcohol-consumption>

1. Questão

Variáveis Qualitativas: Country

Variáveis Quantitativas: beer_servings, wine_servings, spirit_servings, total_litres_of_pure_alcohol

2. Questão

```
library(readxl)
dados <- read.csv("drinks.csv")
```

```
#install.packages('tidyverse')
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.0
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
# Frequência Pontual para variável beer_servings
```

```
beer_servings <- dados$beer_servings
```

```
tabela_freq_pontual_beer_servings <- table(beer_servings)
```

```
print("Frequência pontual para *Porção de Cerveja: ")
```

```
## [1] "Frequência pontual para *Porção de Cerveja: "
```

```
print(tabela_freq_pontual_beer_servings)
```

```
## beer_servings
```

```
## 0 1 2 3 5 6 8 9 12 13 15 16 17 18 19 20 21 22 23 25
## 15 4 1 1 5 4 3 4 1 1 2 2 1 1 2 2 4 1 1 4
## 26 28 31 32 34 36 37 42 43 44 45 47 49 51 52 53 56 57 58 60
## 1 1 4 2 1 3 1 2 1 1 1 1 1 2 3 1 2 1 1 1
## 62 63 64 69 71 76 77 78 79 82 85 88 89 90 92 93 98 99 102 105
## 2 1 1 1 1 2 3 1 1 2 1 1 1 1 1 2 1 1 1 1
## 106 109 111 115 120 122 124 127 128 130 133 140 142 143 144 147 149 152 157 159
## 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1
## 162 163 167 169 171 173 185 188 192 193 194 196 197 199 203 206 213 217 219 224
## 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 1 2
## 225 230 231 233 234 236 238 240 245 247 249 251 261 263 270 279 281 283 284 285
## 1 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 1 1 1
## 295 297 306 313 333 343 346 347 361 376
## 1 1 1 1 1 2 1 1 1 1
```

```
# Frequência Pontual para variável spirit_servings
```

```
spirit_servings <- dados$spirit_servings
```

```
tabela_freq_pontual_spirit_servings <- table(spirit_servings)
```

```
print("Frequência pontual para *Porção de Bebidas Destiladas: ")
```

```
## [1] "Frequência pontual para *Porção de Bebidas Destiladas: "
```

```
print(tabela_freq_pontual_spirit_servings)
```

```
## spirit_servings
```

```
## 0 1 2 3 4 5 6 7 9 11 12 13 15 16 18 19 21 22 25 27
## 23 9 7 8 3 2 3 1 1 2 1 1 2 2 4 1 2 2 2 1
## 29 31 34 35 38 39 41 42 44 46 50 51 55 56 57 60 61 63 65 67
## 1 2 1 4 1 1 2 2 1 1 1 1 1 1 1 1 1 2 1 1
## 68 69 71 72 74 75 76 79 81 84 87 88 97 98 100 101 104 112 114 117
## 1 3 2 1 1 1 2 1 1 1 2 1 2 2 4 1 2 1 3 2
## 118 122 124 126 128 131 132 133 135 137 138 145 147 151 152 154 156 157 158 160
## 2 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1
## 170 173 176 178 179 186 189 192 194 200 202 205 215 216 221 226 237 244 246 252
## 1 2 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1
## 254 258 286 293 302 315 326 373 438
## 1 1 1 1 1 1 2 1 1
```

```
#Frequência Pontual para variável wine_servings
```

```
wine_servings <- dados$wine_servings
```

```
tabela_freq_pontual_wine_servings <- table(wine_servings)
```

```
print("Frequência pontual para *Porção de Vinho: ")
```

```
## [1] "Frequência pontual para *Porção de Vinho: "
```

```
print(tabela_freq_pontual_wine_servings)
```

```
## wine_servings
```

```
## 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 18 19 20 21
## 31 25 9 4 5 6 1 9 7 6 2 5 1 1 1 4 4 1 1 2
## 23 24 26 28 31 32 35 36 42 45 51 54 56 59 62 71 73 74 78 81
```

```
## 1 1 1 1 1 2 1 1 1 3 2 1 2 2 1 1 1 2 1 1
## 84 86 94 97 100 112 113 116 120 123 127 128 129 134 140 149 165 167 172 175
## 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2
## 185 186 190 191 195 212 218 220 221 233 237 254 271 276 278 280 312 339 370
## 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1

# Frequência Pontual para variável total_litres_of_pure_alcohol
total_litres_of_pure_alcohol <- dados$total_litres_of_pure_alcohol
tabela_freq_pontual_total_litres_of_pure_alcohol <- table(total_litres_of_pure_alcohol)

print("Frequência pontual para *Total de litros de Álcool puro: ")

## [1] "Frequência pontual para *Total de litros de Álcool puro: "
print(tabela_freq_pontual_total_litres_of_pure_alcohol)

## total_litres_of_pure_alcohol
## 0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5
## 13 7 4 3 2 3 2 3 1 2 4 3 1 4 1 3
## 1.7 1.8 1.9 2 2.2 2.3 2.4 2.5 2.6 2.8 3 3.1 3.4 3.5 3.8 3.9
## 2 3 1 3 6 2 3 3 2 2 1 1 1 1 2 1
## 4 4.1 4.2 4.3 4.4 4.6 4.7 4.9 5 5.4 5.5 5.6 5.7 5.8 5.9 6.1
## 2 1 4 1 1 2 2 4 1 2 1 1 1 2 3 1
## 6.2 6.3 6.4 6.5 6.6 6.7 6.8 6.9 7 7.1 7.2 7.3 7.6 7.7 8.2 8.3
## 2 5 2 1 4 2 4 1 2 1 3 1 1 2 3 3
## 8.7 8.9 9.1 9.3 9.4 9.5 9.6 9.7 9.8 10 10.1 10.2 10.3 10.4 10.5 10.6
## 1 2 1 1 1 1 1 1 1 2 1 2 1 4 2 1
## 10.9 11 11.3 11.4 11.5 11.8 11.9 12.4 12.9 14.4
## 1 1 2 3 1 2 1 1 1 1

# Frequência Pontual para variável Country
country <- dados$country
tabela_freq_pontual_country <- table(country)

print("Tabela de Frequência Pontual de Países:")

## [1] "Tabela de Frequência Pontual de Países:"
print(tabela_freq_pontual_country)

## country
## Afghanistan Albania
## 1 1
## Algeria Andorra
## 1 1
## Angola Antigua & Barbuda
## 1 1
## Argentina Armenia
## 1 1
## Australia Austria
## 1 1
## Azerbaijan Bahamas
## 1 1
## Bahrain Bangladesh
## 1 1
## Barbados Belarus
## 1 1
```

##	Belgium	Belize
##	1	1
##	Benin	Bhutan
##	1	1
##	Bolivia	Bosnia-Herzegovina
##	1	1
##	Botswana	Brazil
##	1	1
##	Brunei	Bulgaria
##	1	1
##	Burkina Faso	Burundi
##	1	1
##	Cabo Verde	Cambodia
##	1	1
##	Cameroon	Canada
##	1	1
##	Central African Republic	Chad
##	1	1
##	Chile	China
##	1	1
##	Colombia	Comoros
##	1	1
##	Congo	Cook Islands
##	1	1
##	Costa Rica	Cote d'Ivoire
##	1	1
##	Croatia	Cuba
##	1	1
##	Cyprus	Czech Republic
##	1	1
##	Denmark	Djibouti
##	1	1
##	Dominica	Dominican Republic
##	1	1
##	DR Congo	Ecuador
##	1	1
##	Egypt	El Salvador
##	1	1
##	Equatorial Guinea	Eritrea
##	1	1
##	Estonia	Ethiopia
##	1	1
##	Fiji	Finland
##	1	1
##	France	Gabon
##	1	1
##	Gambia	Georgia
##	1	1
##	Germany	Ghana
##	1	1
##	Greece	Grenada
##	1	1
##	Guatemala	Guinea
##	1	1

##	Guinea-Bissau	Guyana
##	1	1
##	Haiti	Honduras
##	1	1
##	Hungary	Iceland
##	1	1
##	India	Indonesia
##	1	1
##	Iran	Iraq
##	1	1
##	Ireland	Israel
##	1	1
##	Italy	Jamaica
##	1	1
##	Japan	Jordan
##	1	1
##	Kazakhstan	Kenya
##	1	1
##	Kiribati	Kuwait
##	1	1
##	Kyrgyzstan	Laos
##	1	1
##	Latvia	Lebanon
##	1	1
##	Lesotho	Liberia
##	1	1
##	Libya	Lithuania
##	1	1
##	Luxembourg	Macedonia
##	1	1
##	Madagascar	Malawi
##	1	1
##	Malaysia	Maldives
##	1	1
##	Mali	Malta
##	1	1
##	Marshall Islands	Mauritania
##	1	1
##	Mauritius	Mexico
##	1	1
##	Micronesia	Moldova
##	1	1
##	Monaco	Mongolia
##	1	1
##	Montenegro	Morocco
##	1	1
##	Mozambique	Myanmar
##	1	1
##	Namibia	Nauru
##	1	1
##	Nepal	Netherlands
##	1	1
##	New Zealand	Nicaragua
##	1	1

##	Niger	Nigeria
##	1	1
##	Niue	North Korea
##	1	1
##	Norway	Oman
##	1	1
##	Pakistan	Palau
##	1	1
##	Panama	Papua New Guinea
##	1	1
##	Paraguay	Peru
##	1	1
##	Philippines	Poland
##	1	1
##	Portugal	Qatar
##	1	1
##	Romania	Russian Federation
##	1	1
##	Rwanda	Samoa
##	1	1
##	San Marino	Sao Tome & Principe
##	1	1
##	Saudi Arabia	Senegal
##	1	1
##	Serbia	Seychelles
##	1	1
##	Sierra Leone	Singapore
##	1	1
##	Slovakia	Slovenia
##	1	1
##	Solomon Islands	Somalia
##	1	1
##	South Africa	South Korea
##	1	1
##	Spain	Sri Lanka
##	1	1
##	St. Kitts & Nevis	St. Lucia
##	1	1
##	St. Vincent & the Grenadines	Sudan
##	1	1
##	Suriname	Swaziland
##	1	1
##	Sweden	Switzerland
##	1	1
##	Syria	Tajikistan
##	1	1
##	Tanzania	Thailand
##	1	1
##	Timor-Leste	Togo
##	1	1
##	Tonga	Trinidad & Tobago
##	1	1
##	Tunisia	Turkey
##	1	1

```
##          Turkmenistan          Tuvalu
##              1              1
##          Uganda          Ukraine
##              1              1
##    United Arab Emirates    United Kingdom
##              1              1
##          Uruguay          USA
##              1              1
##          Uzbekistan          Vanuatu
##              1              1
##          Venezuela          Vietnam
##              1              1
##          Yemen          Zambia
##              1              1
##          Zimbabwe
##              1
```

```
# Tabelas de Frequências Intervalares das variáveis:
```

```
num_intervalos <- 5
tabela_freq_intervalar_beer_servings <- cut(beer_servings, breaks = num_intervalos)
freq_intervalar <- table(tabela_freq_intervalar_beer_servings)
```

```
print("Tabela de Frequência Intervalar para *Porção de Cerveja:")
```

```
## [1] "Tabela de Frequência Intervalar para *Porção de Cerveja:"
```

```
print(freq_intervalar)
```

```
## tabela_freq_intervalar_beer_servings
## (-0.376,75.2]    (75.2,150]    (150,226]    (226,301]    (301,376]
##              96              38              27              23              9
```

```
# Criar tabela de frequência intervalar para spirit_servings
```

```
num_intervalos <- 5
tabela_freq_intervalar_spirit_servings <- cut(spirit_servings, breaks = num_intervalos)
freq_intervalar <- table(tabela_freq_intervalar_spirit_servings)
```

```
print("Tabela de Frequência Intervalar para *Porção de Bebidas Destiladas:")
```

```
## [1] "Tabela de Frequência Intervalar para *Porção de Bebidas Destiladas:"
```

```
print(freq_intervalar)
```

```
## tabela_freq_intervalar_spirit_servings
## (-0.438,87.6]    (87.6,175]    (175,263]    (263,350]    (350,438]
##              120              44              21              6              2
```

```
# Criar tabela de frequência intervalar para wine_servings
```

```
num_intervalos <- 5
tabela_freq_intervalar_wine_servings <- cut(wine_servings, breaks = num_intervalos)
freq_intervalar <- table(tabela_freq_intervalar_wine_servings)
```

```
print("Tabela de Frequência Intervalar para *Porção de Vinho:")
```

```
## [1] "Tabela de Frequência Intervalar para *Porção de Vinho:"
```

```
print(freq_intervalar)
```

```
## tabela_freq_intervalar_wine_servings
```

```
## (-0.37,74] (74,148] (148,222] (222,296] (296,370]
##          150          17          16          7          3

# Criar tabela de frequência intervalar para total_litres_of_pure_alcohol
num_intervalos <- 5
tabela_freq_intervalar_total_litres_of_pure_alcohol <- cut(total_litres_of_pure_alcohol, breaks = num_intervalos)
freq_intervalar <- table(tabela_freq_intervalar_total_litres_of_pure_alcohol)

print("Tabela de Frequência Intervalar para *Total de litros de Álcool puro:")

## [1] "Tabela de Frequência Intervalar para *Total de litros de Álcool puro:"
print(freq_intervalar)

## tabela_freq_intervalar_total_litres_of_pure_alcohol
## (-0.0144,2.88] (2.88,5.76] (5.76,8.64] (8.64,11.5] (11.5,14.4]
##          83          30          43          31          6

## 3. Questão
library(psych)

##
## Attaching package: 'psych'

## The following objects are masked from 'package:ggplot2':
##
##      %+%, alpha

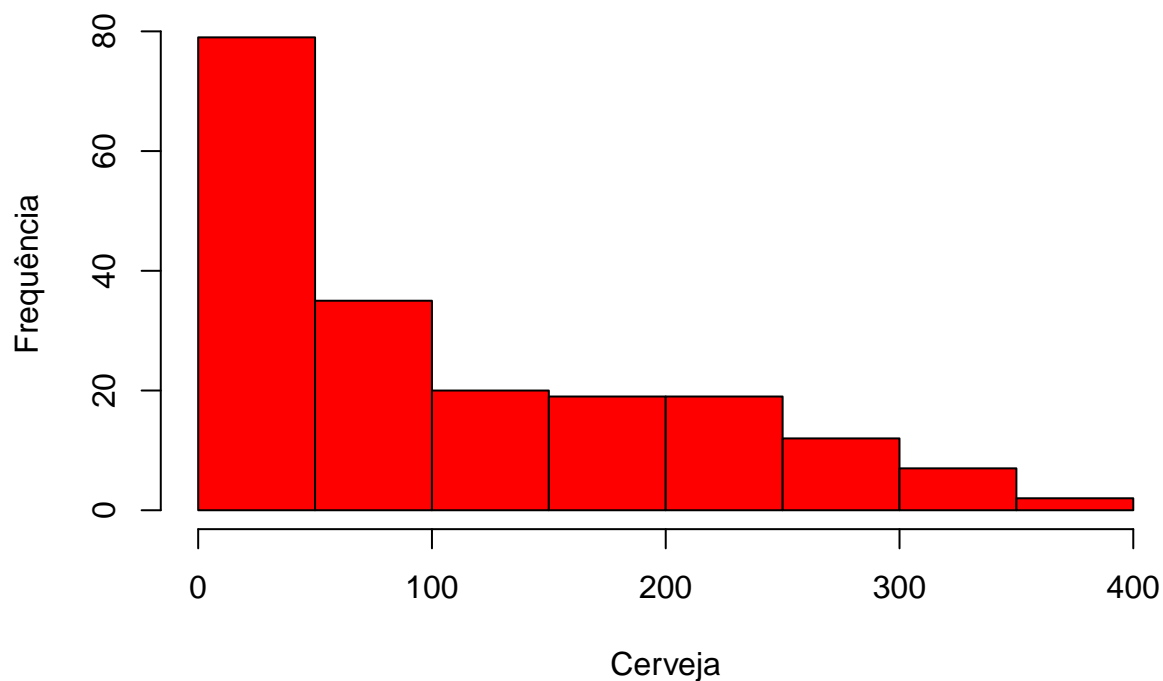
# Calcular medidas-resumo
resumo = describe(dados)
print(resumo)

##              vars    n  mean    sd median trimmed  mad min
## country*         1 193  97.00  55.86   97.0   97.00 71.16   1
## beer_servings     2 193 106.16 101.14   76.0   94.30 99.33   0
## spirit_servings    3 193  80.99  88.28   56.0   67.03 80.06   0
## wine_servings     4 193  49.45  79.70    8.0   31.28 11.86   0
## total_litres_of_pure_alcohol 5 193   4.72   3.77    4.2    4.46  4.45   0
##              max range skew kurtosis   se
## country*    193.0 192.0 0.00   -1.22 4.02
## beer_servings 376.0 376.0 0.80   -0.51 7.28
## spirit_servings 438.0 438.0 1.27    1.36 6.35
## wine_servings 370.0 370.0 1.88    2.78 5.74
## total_litres_of_pure_alcohol 14.4 14.4 0.42   -1.01 0.27

## 4. Questão

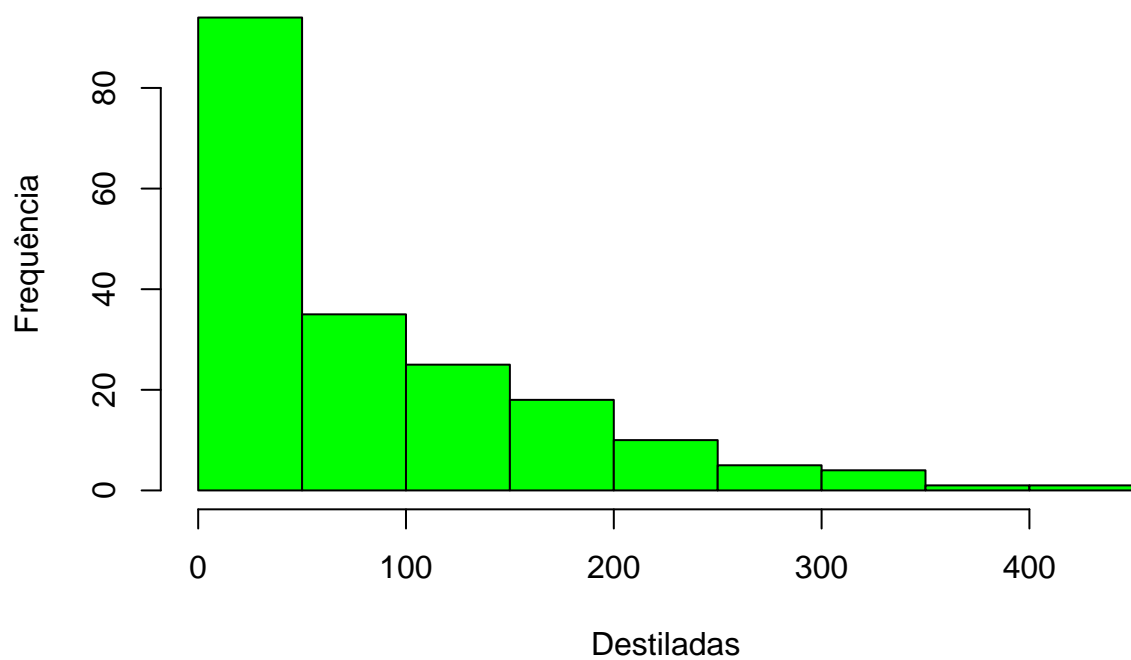
#Construir o Histograma de cada variável quantitativa.
hist(beer_servings, main = "Histograma das Cervejas", xlab = "Cerveja", ylab = "Frequência", col="red",
```


Histograma das Cervejas



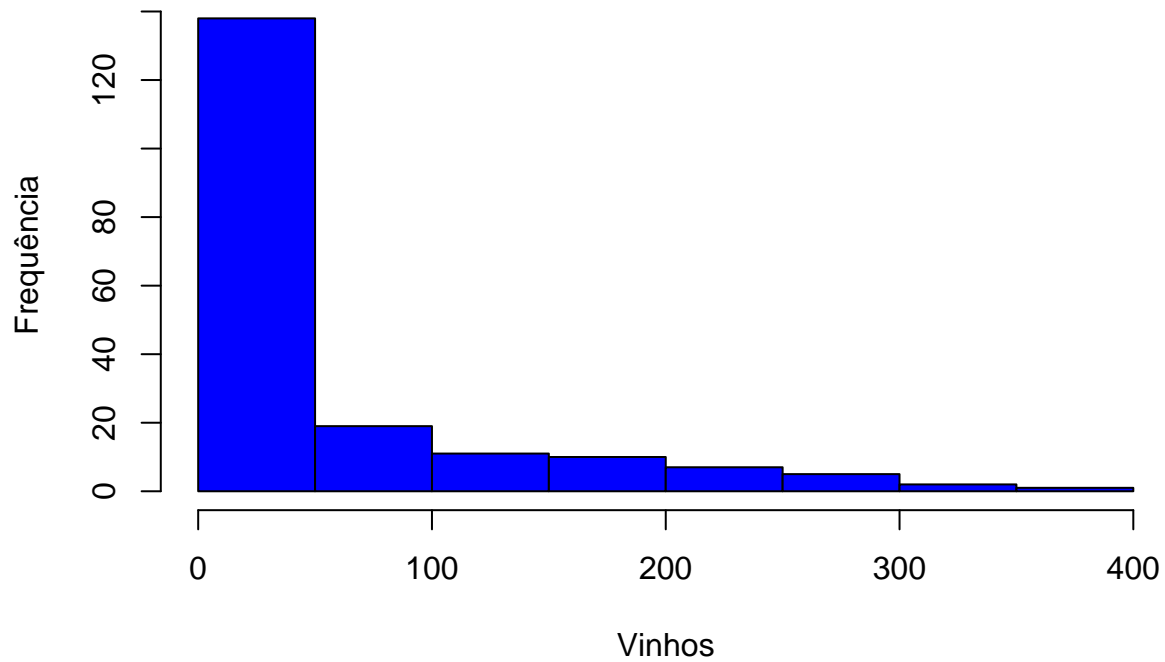
```
hist(spirit_servings, main = "Histograma das Bebidas Destiladas", xlab = "Destiladas", ylab = "Frequência", col="red", border="black")
```

Histograma das Bebidas Destiladas



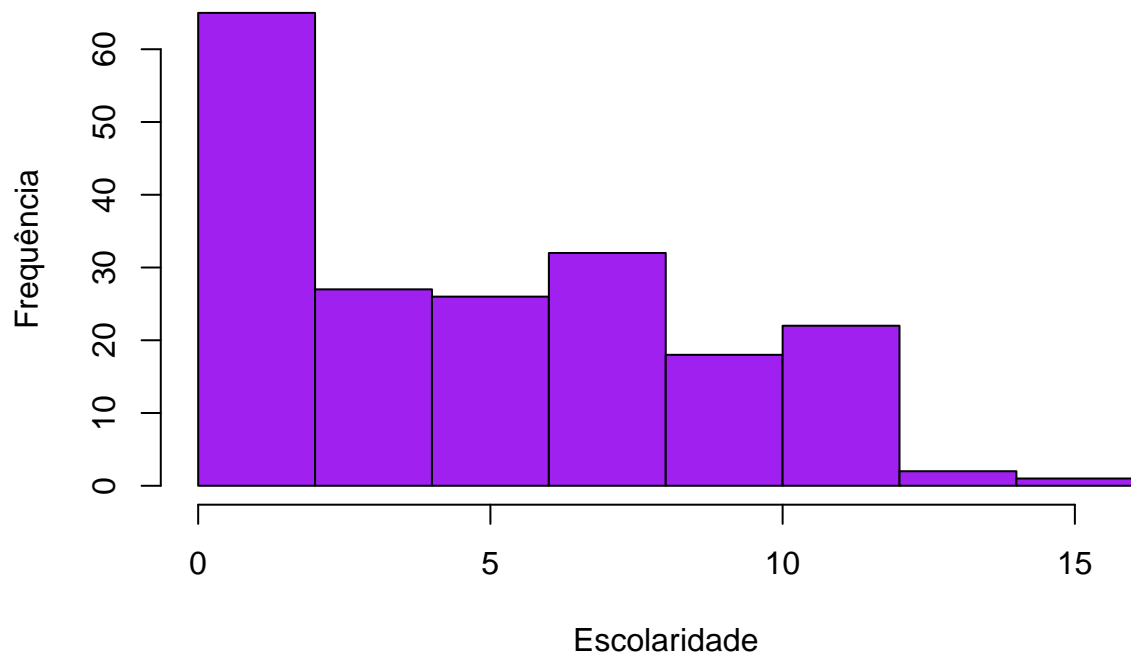
```
hist(wine_servings, main = "Histograma de Vinhos", xlab = "Vinhos", ylab = "Frequência", col="blue", border="black")
```

Histograma de Vinhos



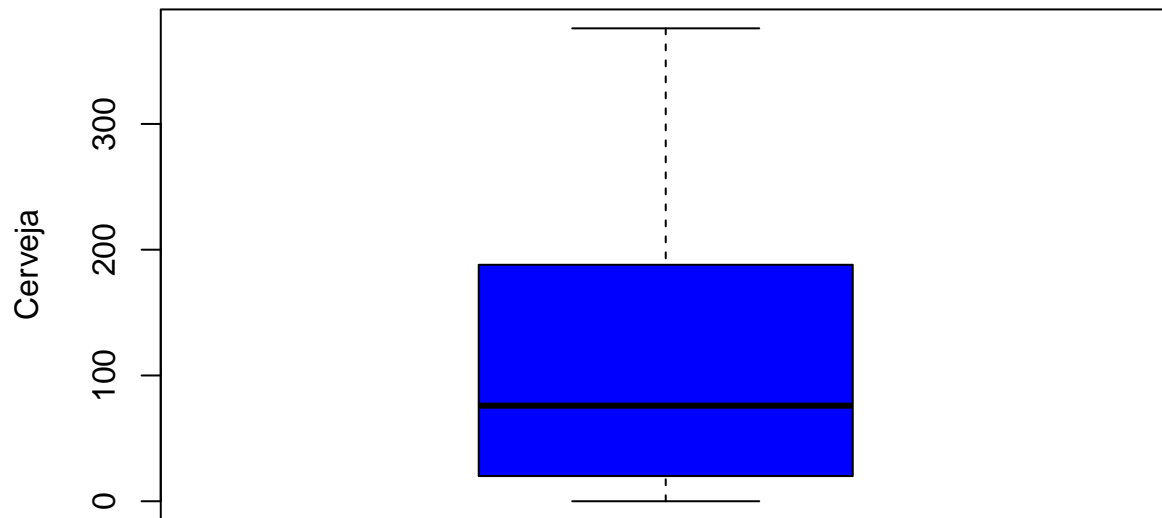
```
hist(total_litres_of_pure_alcohol, main = "Histograma de Álcool puro", xlab = "Escolaridade", ylab = "F
```

Histograma de Álcool puro



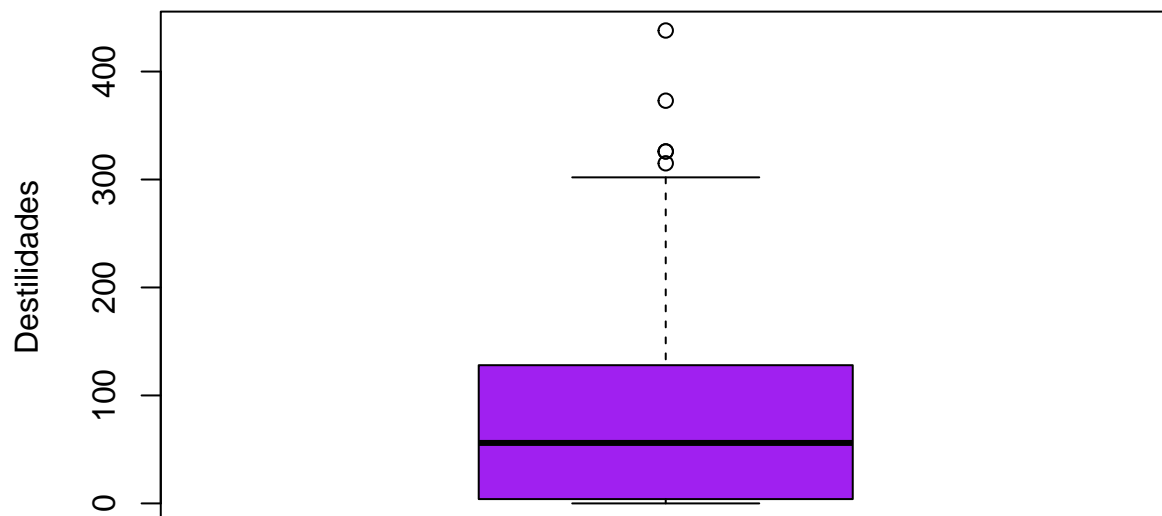
```
#Construir o Boxplot de cada variável quantitativa.  
boxplot(beer_servings, main = "Boxplot das Cervejas", ylab = "Cerveja", col="blue", border="black")
```

Boxplot das Cervejas



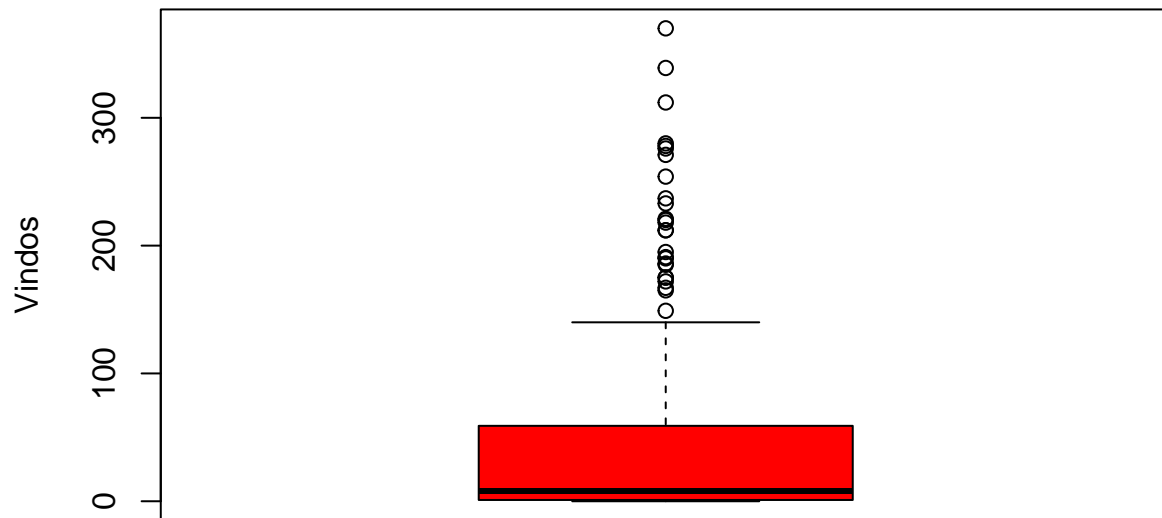
```
boxplot(spirit_servings, main = "Boxplot das Bebidas Destilidades", ylab = "Destilidades", col="purple")
```

Boxplot das Bebidas Destiladas



```
boxplot(wine_servings, main = "Boxplot dos Vinhos", ylab = "Vinhos", col="red", border="black")
```

Boxplot dos Vinhos



```
boxplot(total_litres_of_pure_alcohol, main = "Boxplot dos Litros de Puro Álcool", ylab = "Puro Álcool",
```

Boxplot dos Litros de Puro Álcool

