

ME115 - Linguagem R

Atividade Prática 01

1º semestre de 2023

Introdução

Operações exploradas:

- criação de vetores utilizando: `c()`, `:`, `seq()` e `rep()`;
- indexação (`[]`) de vetores através de: índice numérico, vetores numéricos, classes dos vetores, índices resultantes de operações lógicas (usamos o `<`);
- classe de objeto (comando `class()`);
- criação e exploração do objeto `matrix`;
- instalação (`install.packages()`) e carregamento (`library()`) de pacotes;
- criação e exploração do objeto `data.frame`;
- realização de gráfico usando o comando `plot()` usando escala original e logarítmica das variáveis.

Exercícios

Os exercícios abaixo são do Capítulo 2 do Livro: <https://rafalab.github.io/dsbook/r-basics.html>

1. Qual é a soma dos 100 primeiros números inteiros positivos? Como você faria isso no R?
2. Agora faça a mesma coisa para calcular a soma dos números inteiros de 1 a 1000.
3. Observe o resultado ao digitar o seguinte código em R:

```
n <- 1000
x <- seq(1, n)
sum(x)
```

Com base no resultado, o que você acha que as funções `seq` e `sum` fazem? Explore a ajuda do R.

4. Use a função `c()` para criar um vetor com as médias das temperaturas máximas em janeiro nas cidades: Beijing, Lagos, Paris, Rio de Janeiro, San Juan e Toronto, que são, respectivamente, 35, 88, 42, 84, 81, and 30 graus Fahrenheit. Chame este objeto de `temp`.
5. Crie um vetor com os nomes das cidades e chame esse objeto de `city`.
6. Use a função `names()` e os objetos definidos nos itens anteriores para associar às temperaturas os nomes das cidades correspondentes.
7. Considerando o vetor `temp` que você criou:
 - a) Use os operadores `[` e `:` para acessar as temperaturas das três primeiras cidades listadas.
 - b) Use o operador `[` operator para acessar as temperaturas de Paris e San Juan.
8. Use o operador `:` para criar a sequência de números: 12, 13, 14, ..., 73.
9. Crie um vetor contendo todos os números ímpares positivos e menores do que 100.
10. Crie um vetor de números que inicia em 6, não ultrapassa 55 e aumenta em incrementos de 4/7: 6, 6 + 4/7, 6 + 8/7, etc. Quantos elementos estão neste vetor? Dica: estude a função `seq()` e seus argumentos.

11. Qual é a classe dos seguintes objetos: `a <- seq(1, 10, 0.5)` e `b <- seq(1, 10)`?
12. Defina o seguinte vetor: `x <- c("1", "3", "5")` e force que ele contenha números inteiros.
13. Crie a seguinte matriz:

$$A_{4 \times 5} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 & 101 & 105 \\ 2 & 6 & 10 & 102 & 106 \\ 3 & 7 & 11 & 103 & 107 \\ 4 & 8 & 12 & 104 & 108 \end{bmatrix}.$$

14. Considerando a matriz que você criou no Exercício 13:
 - a) Use os operadores `[` e `:` para acessar os elementos a_{21} , a_{22} , a_{23} .
 - b) Acesse os elementos da terceira coluna.
 - c) Acesse os elementos da quarta linha.
 - d) Altere o valor a_{11} por 20.
15. Você pode criar um conjunto de dados usando a função `data.frame`. Exemplo:

```
temp <- c(35, 88, 42, 84, 81, 30)
city <- c("Beijing", "Lagos", "Paris", "Rio de Janeiro",
         "San Juan", "Toronto")
city_temps <- data.frame(name = city, temperature = temp)
```

Use o código acima para criar o conjunto de dados, mas adicione uma variável que indique a temperatura em Celsius.

16. Carregue o pacote `dslabs`. Veja a ajuda desse pacote. Carregue também o conjunto de dados `murders`. Qual a classe de `murders`?
17. Use a função `str()` para explorar a estrutura do conjunto de dados `murders`.

Leia a seção [2.13] - Indexação (<https://rafalab.github.io/dsbook/r-basics.html#indexing>)

Calcule a taxa de mortes por arma de fogo por 100 mil habitantes para cada estado e guarde em um objeto com nome `murder_rate`. Use um operador lógico para criar um vetor lógico chamado `low` que indica quais estados possuem taxa abaixo de 1.

Extras

Os exercícios 18 a 25 não foram feitos em aula, mas vocês podem tentar resolvê-los.

18. Use os resultados do exercício anterior e a função `which()` para determinar os índices (posições do vetor) de `murder_rate` que estão abaixo de 1.
19. Use os resultados do exercício anterior para listar os nomes dos estados com taxas de mortes por arma de fogo abaixo de 1 (por 100 mil habitantes).
20. Liste os estados da região “Northeast” com taxas de mortes por arma de fogo abaixo de 1 (por 100 mil habitantes). Dica: use o vetor lógico `low` e o operador lógico `&`.
21. Calcule a média das taxas de mortes por arma de fogo por 100 mil habitantes entre os estados.
22. Quantos estados estão abaixo da média?

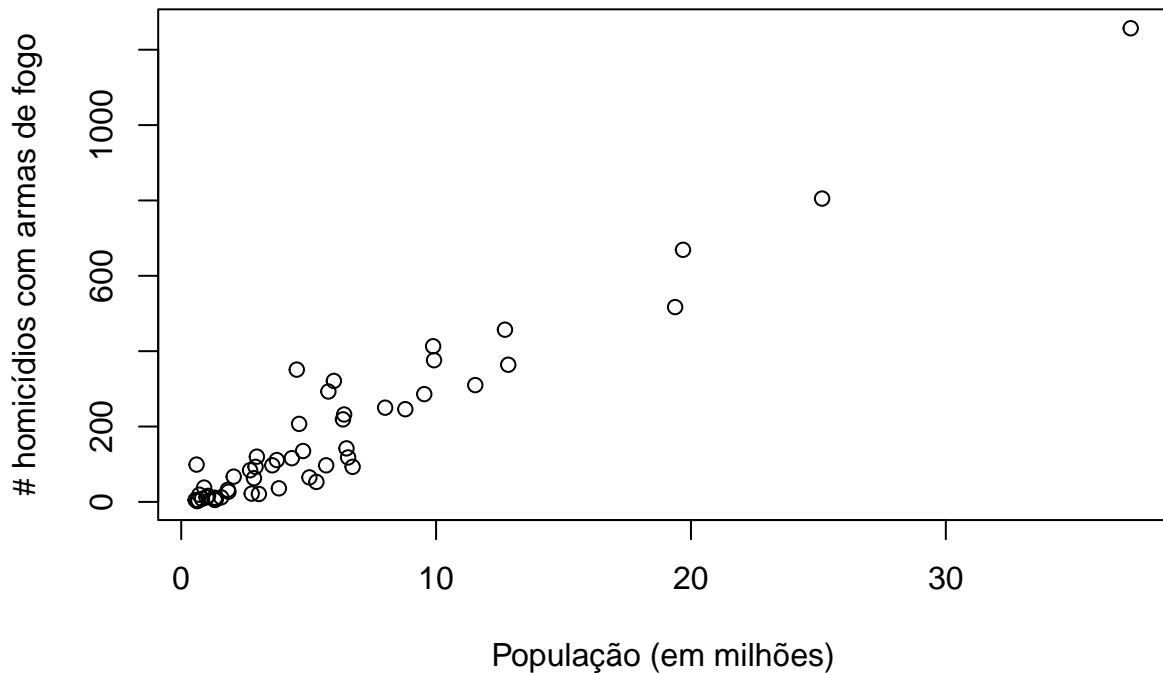
Leia a seção [2.15] - Gráficos básicos (<https://rafalab.github.io/dsbook/r-basics.html#basic-plots>)

23. O código a seguir produz um gráfico do total de mortes versus tamanho da população.

```
library(dslabs)
data(murders)

population_in_millions <- murders$population/10^6
total_gun_murders <- murders$total

plot(population_in_millions, total_gun_murders,
      ylab="# homicídios com armas de fogo",
      xlab="População (em milhões)")
```



Como muitos estados têm população abaixo de 5 milhões, o gráfico não mostra muito bem os dados. Refaça o gráfico usando a escala logarítmica. Transforme usando `log10()` e refaça o gráfico.

24. Apresente um histograma do total da população de cada estado. Dica: `hist()`.
25. Apresente boxplots do total da população de cada estado por região (cada boxplot representa uma região diferente).

Agradecimentos

Esse material foi incrementado do material produzido pelas Profas. Samara Kiihl e Tatiana Benaglia para o curso de ME115.