



## Programação em R - Homework

tembici

A lista compreende exercícios sobre programação em R. Utilize os materiais abaixo como base para a solução:

- [Programação em R - Introdução](#)
- [Programação em R - Intermediário](#)

**VETORES:** Vetores são uma estrutura de dados básica do R, que permite armazenar um conjunto de valores numéricos, caracteres ou lógicos sob um mesmo nome. Esses valores podem ser acessados individualmente através de índices. Abaixo, você encontrará diversos exercícios sobre vetores. Faça uso do seu R/RStudio para encontrar as respostas.

1. Considere um vetor  $x = c(4, 6, 5, 7, 10, 9, 4, 15)$ . Qual é o valor de  $x < 7$ ?
2. Considere os vetores  $p = c(3, 5, 6, 8)$  e  $q = c(3, 3, 3)$ . Qual é o resultado de  $p+q$ ?
3. Considere o vetor  $x = c(1, 2, 3, 4)$ . Qual é o valor de  $k$  para  $(x+2)[(!is.na(x)) \& x > 0] \rightarrow k$ ?
4. Se  $x = c(2, 4, 6, 8)$  e  $y = c(TRUE, TRUE, FALSE, TRUE)$ . Qual é valor de  $sum(x[y])$ ?
5. Considere o vetor  $x = c(34, 56, 55, 87, NA, 4, 77, NA, 21, NA, 39)$ . Qual é o código que contará a quantidade de NAs em  $x$ ?
6. Considere os vetores  $x = c(4, 6, 5, 7, 10, 9, 4, 15)$  e  $y = c(0, 10, 1, 8, 2, 3, 4, 1)$ . Qual é o valor de  $x*y$ ?
7. Considere os vetores  $a = c(1, 2, 4, 5, 6)$  e  $b = c(3, 2, 4, 1, 9)$ . Qual é o valor de  $cbind(a, b)$ ?
8. Considere os vetores  $a = c(1, 5, 4, 3, 6)$  e  $b = c(3, 5, 2, 1, 9)$ . Qual é o valor de  $a <= b$ ?

**FATORES:** Os fatores são uma classe especial de vetores, que definem variáveis categóricas. Nesta seção você entenderá como executar operações com fatores e declarar variáveis como fatores.

1. Se  $x = c(1, 2, 3, 3, 5, 3, 2, 4, NA)$ , quais são os *levels* de  $factor(x)$ ?
2. Deixe  $x = c(11, 22, 47, 47, 11, 47, 11)$ . Se  $factor(x, levels=c(11, 22, 47), ordered=TRUE)$  é executado, qual será o quarto elemento do resultado?
3. Se  $s1 = factor(sample(letters, size=5, replace=TRUE))$  e  $s2 = factor(sample(letters, size=5, replace=TRUE))$ , escreva o código em R que concatenará  $s1$  e  $s2$  em um único fator com 10 elementos.
4. Se  $x <- factor(c("high", "low", "medium", "high", "high", "low", "medium"))$ , escreva o código em R que fornecerá a quantidade de observações para cada *level* de  $x$ .

**MATRIZES:** Matrizes, são tabelas retangulares de números, expressões matemáticas ou símbolos, cujo elementos são arranjados em  $n$  linhas e  $p$  colunas.

1. Seja  $M = matrix(c(1:10), nrow=5, ncol=2, dimnames=list(c('a', 'b', 'c', 'd', 'e'), c('A', 'B')))$ . Encontre:
  - $M[1,];$
  - $M[,1];$
  - $M[3,2];$
  - $M['e', 'A']$
2. Considere  $A = matrix(c(2, 0, 1, 3), ncol=2)$  e  $B = matrix(c(5, 2, 4, -1), ncol=2)$ . Encontre:

- $A + B$
  - $A - B$
3. Encontre a solução para  $Ab$ , onde  $A$  é a matriz definida no exercício anterior e  $b = c(7, 4)$ .
  4. Qual é a transposta de  $A$ ?
  5. Qual é a inversa de  $A$ ?
  6. Encontre o determinante de  $B$ .

**DATA FRAMES:** Até agora temos utilizado apenas dados de uma mesma classe, armazenados ou em um vetor ou em uma matriz. Mas uma base de dados, em geral, é feita de dados de diversas classes diferentes. Por exemplo, podemos querer ter uma coluna com os nomes dos funcionários, outra com o sexo dos funcionários, outra com valores... note que essas colunas são de classes diferentes, como textos e números. Como guardar essas informações?

A solução para isso é o `data.frame`. O `data.frame` é talvez o formato de dados mais importante do R. No `data.frame` cada coluna representa uma variável e cada linha uma observação. Essa é a estrutura ideal para quando você tem várias variáveis de classes diferentes em um banco de dados.

1. Crie o seguinte data frame no R. Não esqueça de definir a variável **Sex** como fator.

	Age	Height	Weight	Sex
Alex	25	177	57	M
Lilly	31	163	69	M
Mark	23	190	83	F
Oliver	52	179	75	F
Martha	76	163	70	M
Lucas	49	183	83	F
Caroline	26	164	53	M

2. Crie o data frame abaixo (aqui, a variável **Working** é um *character* e não um *factor*).

	Working
Alex	Yes
Lilly	No
Mark	No
Oliver	Yes
Martha	Yes
Lucas	No
Caroline	Yes

3. Adicione o data frame de apenas uma coluna do exercício 2 no data frame encontrado em 1.
4. Quantas linhas e colunas tem o novo data frame encontrado em 3?
5. Qual é a classe de cada coluna do data frame encontrado em 3?

**LISTAS:** A lista no R nos permite armazenar uma variedade de objetos (matrizes, vetores, `data.frames` ou até mesmo outra lista) e tais objetos podem não estar relacionados uns com os outros. Desta forma, uma lista tem grande utilidade quando queremos armazenar qualquer tipo de dado em um único objeto sem nos preocupar com o tamanho dos objetos envolvidos.

1. Seja  $p = c(2, 7, 8)$ ,  $q = c("A", "B", "C")$  e  $x = list(p, q)$ . Qual o valor de  $x[2]$ ?
2. Seja  $w = c(2, 7, 8)$ ,  $v = c("A", "B", "C")$  e  $x = list(w, v)$ . Qual é o código em R que substituirá "A" em  $x$  por "K"?
3. Se  $a = list("x"=5, "y"=10, "z"=15)$ , qual o código em R que retornará a soma dos elementos em  $a$ ?
4. Considere 3 vectors:  $year=c(2005:2016)$ ,  $month=c(1:12)$  e  $day=c(1:31)$ . Defina uma lista que tenha como resultado o seguinte:

```
# $year
# [1] 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016
```

```
#
# $month
# [1] 1 2 3 4 5 6 7 8 9 10 11 12
#
# $day
# [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
# [24] 24 25 26 27 28 29 30 31
```

**CONDIÇÕES:** Comumente temos que fazer escolhas dada uma condição. Se alguma coisa for verdade, então iremos agir de uma maneira, caso contrário, então podemos tomarmos outra ação. Condições são úteis em análise de dados.

1. Crie um código em R que retorne o valor absoluto de um número inteiro.
2. Crie um código em R que calcula a raiz quadrada de um dado valor inteiro, se o valor for negativo o código deveria retornar `NA`.
3. Crie um código em R que retorna o valor máximo de um vetor de números de tamanho 2.
4. Crie um código em R que retorna a quantidade de valores que são maiores que a média de um vetor. Você pode usar a função `mean()`.
5. Crie um código em R que, dado um vetor numérico de tamanho 3, mostrará os elementos ordenados do maior para o menor.
6. Qual é o resultado de `ifelse(sqrt(9)<2,sqrt(9),0)`?
7. Qual é o valor de `x` e `y` para: `z=-1` e `if(z<0){x=abs(z);y=z*3}`?

**LOOPS:** Muitas vezes pode ser preciso reescrever o seu código em função da necessidade de se repetir uma atividade. Isso pode ser trabalhoso e consumir bastante tempo. Felizmente, temos em programação a opção de usar `loops` (no R, as funções `for()`, `while()` e `repeat()`).

1. Usando `msg = c("Hello")` e `i = 1`, escreva um loop no formato `while()` que incrementa a variável `i` 6 vezes e mostra `msg` em cada iteração (neste caso, `while(i<7)`)
2. Escreva um loop no formato `for()` que mostra os primeiros quatro números da sequência: `x <- c(7, 4, 3, 8, 9, 25)`.
3. Escreva um loop no formato `for()` que mostra todos os elementos em `y <- c("q", "w", "e", "r", "z", "c")`.

## FUNÇÕES:

1. Crie uma função que retornará a soma de dois números inteiros.
2. Crie uma função que retornará `TRUE` se um dado valor faz parte de um vetor de números.
3. Crie uma função que dado um vetor e um número inteiro retornará quantas vezes o inteiro aparece no vetor.