

Introdução à Análise de Dados – Completar para a G1 – 2023.1

1) Manipulação de vetores.

- a) Defina um vetor **v** com a sequência de números ímpares maiores que zero e menores que 50.
- b) Calcule a soma dos elementos desse vetor.
- c) Selecione os elementos deste vetor que são múltiplos de 3 e os armazene em uma variável chamada **v.selecionado**.
- d) Apresente um vetor chamado **v.dividido** que contém a divisão por elemento da sequência obtida no item a) por 3.

2) Utilizando apenas as funções **c()**, **seq()** e **indexação de vetores []**, crie os seguintes objetos:

- a) Uma sequência de vinte valores em intervalos regulares, indo de 0 a 100, nomeada **sq1**;
- b) Um objeto, denominado **sq2**, que contenha todos os elementos de **sq1**, exceto o quinto e décimo quinto valores;
- c) Um vetor **sq3** contendo apenas as posições ímpares do objeto **sq1**;
- d) Uma sequência igual a **sq1** substituindo, apenas, os valores nas posições pares, pelo número relativo à sua posição. Denomine esse objeto de **sq4**.
- e) Utilizando o comando **matrix**, crie uma matriz 3 x 3 chamada **m** a partir dos dados armazenados em **sq1**.
- f) Divida os elementos da matriz **m** por 2.

3) Considere os vetores **u**, **v**, **w**, **x**, **y** e **z** e a matriz **m** formada a partir da união desses vetores, onde **u** será a primeira linha, **v**, a segunda e assim por diante:

$u = (-1, 3, -8, 9, 10)$, $v = (-8, 12, 4, 2, 3)$, $w = (-2, -5, 3, 6, 7)$, $x = (1, 5, 9, 12, 6)$, $y = (2, 6, 10, -3, 8)$, $z = (3, 7, 11, -10, 2)$

Para os vetores **u**, **v**, **w**, **x**, **y** e **z** e a matriz **m**, escreva comandos em R para retornar:

- a) A média dos valores de $x + y - u$.
- b) Um vetor chamado **maior.igual** que identifica como **TRUE** os elementos do vetor resultante de $x + y + z$ cujos valores são maiores ou iguais a 3 e como **FALSE** os demais elementos.
- c) A matriz **m** criada a partir dos vetores.
- d) O elemento que está na linha 3 e coluna 2 da matriz **m**.
- e) Um vetor contendo a segunda coluna da matriz **m**.
- f) Uma matriz chamada **matriz.selecionada** contendo a primeira e a terceira linhas da matriz **m**.

4) No estudo de escalas de temperatura, aprendemos que a conversão de graus Fahrenheit (F) para graus Celsius (C) se dá pela fórmula $F = (9 \cdot C / 5) + 32$.

Elabore uma função em R chamada `converte.temperatura` que receba como parâmetro a temperatura que deve ser convertida e a escala (F para Fahrenheit ou C para Celsius) na qual a mesma se encontra e retorne a temperatura convertida, caso receba uma escala diferente de "F" ou "C" retorna **NaN**.

5) Elabore uma função chamada **conta.primos** que recebe os parâmetros **inicio** e **fim** e retorna o total de números primos no intervalo entre eles. Observe que **inicio** e **fim** também devem ser considerados.

6) Usando as instruções **for** e **if** e as funções **nrow()** e **ncol()**, construa uma função chamada **calcula.soma** que recebe uma matriz **m** e calcula a soma dos elementos da matriz que possuem valor superior a 5.

7) Manipulação de matriz.

- a) Crie um objeto da classe **matriz** chamado **matriz.normal** com 30 linhas e 50 colunas contendo uma amostra de uma distribuição normal de média 10 e desvio padrão 3,6.
- b) Apresente a linha 2 e a seguir a coluna 3 da matriz.
- c) Apresente uma matriz chamada **matriz.selecionada** com os elementos da **matriz.normal** que estão nas posições definidas pelas linhas de 3 a 5 e pelas colunas de 8 a 10.
- d) Apresente as dimensões da matriz.
- e) Calcule a soma dos elementos da matriz.
- f) Calcule o produto da matriz por sua matriz transposta.
- g) Calcule a soma dos elementos da primeira linha da matriz.
- h) Calcule média e variância por linha. Guarde os resultados em um data frame chamado **dataframe.linha**, cujas colunas são a média e a variância por linha e tem o nome de **media** e **variancia**, respectivamente. Faça o mesmo para as colunas, com o data frame de nome **dataframe.coluna** e os nomes das colunas **media** e **variancia**.

8) Manipulação de arrays.

- a) Crie um **array** tri-dimensional com as dimensões 4, 5 e 3 e o preencha com os sessenta primeiros números inteiros.
- b) Calcule a soma dos elementos do **array** cuja coordenada na segunda dimensão é 4.
- c) Calcule a média dos elementos do **array** cuja coordenada nas primeiras duas dimensões é 1.
- d) Calcule o **array** que se obtém multiplicando todos os valores do **array** por 2 e somando 5.

9) Escreva uma função chamada **gera.fibonacci** que retorna um vetor com a sequência Fibonacci dado um parâmetro **n** de entrada que determina o número de elementos da sequência a serem retornados. Use o comando **for** na implementação.

10) Usando a instrução **for** e a função **length()**, construa um procedimento chamado **apresenta.palavra** que recebe um vetor de caracteres (**vetor.caracteres**) e um vetor numérico (**vetor.numerico**). O vetor numérico armazena os índices a serem utilizados pelo procedimento para selecionar e apresentar os caracteres armazenados no vetor de caracteres.

11) Usando a instrução **for**, construa uma função chamada **transforma.vetor** que recebe um vetor de **n*m** posições, e gera como resultado uma matriz **n x m**. Observe que a matriz deve ser preenchida da esquerda para a direita e de cima para baixo, seguindo a ordem de posição dos elementos do vetor. Assim, o primeiro elemento deve estar na posição (1, 1), o segundo, na posição (1, 2) e assim por diante.

12) Usando as instruções **while** e **if** e a função **length()**, construa um procedimento chamado **encontra.valor** que recebe dois vetores numéricos **x** e **y** e apresenta na tela o total de elementos do primeiro vetor que também estão no segundo vetor. Considere também que: se um mesmo elemento de **x** aparecer mais de uma vez em **y**, deverá ser contado como uma única vez; se **x** apresentar elementos repetidos, eles deverão ser considerados como elementos distintos.

13) Usando as instruções **for** e **if** e as funções **nrow()** e **ncol()**, construa uma função chamada **calcula.soma** que recebe uma matriz **m** e calcula a soma dos elementos da matriz que possuem valor superior a **5**.

14) Usando as instruções **for** e **if** e as funções **nrow()** e **ncol()**, construa uma função chamada **calcula.transposta** que recebe uma matriz **m** e calcula a sua matriz transposta. Verifique, usando funções prontas do R, se a função **calcula.transposta** está correta.

15) Usando as instruções **for** e **if** e as funções **nrow()** e **ncol()**, construa uma função chamada **calcula.valor** que recebe uma matriz **m** e calcula o quadrado da soma dos elementos da matriz que possuem valor superior a **3** e inferior a **6**.

16) Uma loja que vende motos deseja comparar vendas e preços de 2018 com as vendas e preços de 2010 de determinadas motos. As motos escolhidas, e seus respectivos preços e quantidades vendidas, estão na tabela a seguir.

Moto	2010		2018	
	Preço (R\$)	Quantidade	Preço (R\$)	Quantidade
Harley Iron	36.000,00	5	45.000,00	4
Honda Titan	12.000,00	10	13.000,00	15
Ninja MT	22.000,00	15	24.000,00	12

Para os dados acima:

- Crie um data frame chamado **loja** contendo cinco colunas a saber: **Moto**, com os modelos das motos; **P2010**, com os preços das motos para 2010; **Q2010**, com a quantidade de motos vendidas em 2010; **P2018**, com os preços das motos para 2018; **Q2018**, com a quantidade de motos vendidas em 2018.
- Use uma função do R para apresentar a estrutura do data frame, e outra para apresentar o nome das colunas.
- Usando funções do R, calcule para **2010** o menor valor para uma moto.
- Selecione no data frame apenas as colunas **P2018** e **Q2018** e respectivas linhas cuja quantidade vendida em **2010** seja superior a **5**. Devem ser utilizados mecanismos ou funções de filtro do R.
- Usando funções do R, crie uma coluna chamada **TotalVenda** no data frame que contenha, para cada uma das linhas do data frame, a soma das quantidades vendidas por moto em **2010** e **2018**.