

Descoberta do Conhecimento – 1ª. Lista

Os Exercícios a seguir devem ser implementados utilizando a Linguagem R. Devem ser entregues o código fonte comentado.

- 1) Escreva uma função chamada `menor.dctres` que dados três números como parâmetros (`val1`, `val2` e `val3`), retorna o menor deles.
- 2) Escreva uma função chamada `calculadora` que receba como parâmetros dois valores numéricos (`val1`, `val2`) e também uma *string* (`oper`) identificando a operação aritmética ("`+`", "`-`", "`*`" ou "`/`") a ser realizada entre os dois parâmetros numéricos e retorne o resultado da respectiva operação.
- 3) Crie uma função chamada `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.
- 4) Manipulação de vetores.
 - a) Defina uma variável com a sequência de números ímpares maiores que zero e menores que 50.
 - b) Calcule a soma dos elementos deste vetor.
 - c) Selecione os membros desta sequência que são múltiplos de 3.
 - d) Apresente um vetor que contém a divisão por elemento da sequência obtida no item anterior por 3.
- 5) 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`.
- 6) As leituras mensais do medidor de consumo de eletricidade de uma casa foram:

Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
8839	9159	9476	9736	10249	10664	11057	11569	11969	12310	12672	13002

- Crie um objeto chamado `eletricidade` com os valores das leituras, de janeiro a dezembro.
 - Calcule o consumo mensal neste período com a função `diff` e guarde o resultado em um objeto chamado `eletricidade.consumo`.
 - Usando o item b), calcule o máximo e o mínimo de consumo mensal e guarde os resultados em um objeto chamado `eletricidade.range`.
 - Calcule a média, mediana, variância e desvio padrão dos consumos mensais de eletricidade e guarde-os em objetos chamados `eletricidade.media`, `eletricidade.mediana`, `eletricidade.variancia` e `eletricidade.desviopadrao`, respectivamente.
- 7) Implemente uma função chamada `desvio.padrao` que recebe como parâmetro uma amostra e fornece como retorno o seu desvio padrão.

$$DP = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2}$$

DP = desvio padrão;

x_i = valor da observação i ;

\bar{x} = média das observações;

N = número de observações.

- 8) Manipulação de matriz.
- Crie um objeto da classe `matriz` chamado `matriz.normal` com 3 linhas e 5 colunas contendo uma amostra de uma distribuição normal de média 10 e desvio padrão 3,6.
 - Apresente a linha 2 e a seguir a coluna 3 da matriz.
 - Apresente as dimensões da matriz.
 - Calcule a soma dos elementos da matriz.
 - Calcule o produto da matriz por sua matriz transposta.
 - Calcule a soma dos elementos da primeira linha da matriz.
 - 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`.

9) 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.

10) Considere a informação constante no arquivo “PopulacaoEconomicamenteAtiva.csv” e importe-os para um data frame chamado `dfPopulacaoEconomicamenteAtiva`.

- a) Apresente as informações de dimensão e nomes das colunas.
- b) A partir do data frame anterior, crie um novo data frame que contenha as informações posteriores a 2003 e referentes somente aos estados RJ, SP, MG e ES.
- c) Referente aos dados obtidos no item anterior, crie um sumário de dados que apresente, para cada ano, o número total de pessoas economicamente ativas.
- d) Apresente um gráfico de barras na cor azul com o sumário de dados anual obtido no item anterior. O eixo x deve ter o rótulo Ano e o eixo y o rótulo “População Economicamente Ativa”.

11) Uma empresa que vende laranja e produz suco de laranja possui um equipamento que coleta informações sobre as laranjas enquanto elas passam pela esteira da fábrica. Este equipamento coleta as seguintes informações: índice de acidez (variando de 0 a 1); peso; diâmetro. Considerando arquivo “Laranjas.csv” que apresenta o resultado gerado por este equipamento após ele analisar 100 laranjas faça os itens a seguir.

- a) Monte um data frame chamado `dfLaranja` com as informações do arquivo “Laranjas.csv”;
- b) Sabendo que o índice de acidez (*i*) da laranja é utilizado para definir sua categoria, monte uma função chamada `define.categoria` que categorize as laranjas e inclua no data frame `dfLaranja` a informação de categoria. As categorias são: A ($i < 0,5$); B ($0,5 < i \leq 0,7$); C ($0,7 < i \leq 0,9$); D ($0,9 < i \leq 1,1$); E ($1,1 < i \leq 1,3$); F ($1,3 < i \leq 1,5$); G ($1,5 < i \leq 1,7$); H ($i > 1,7$).
- c) Crie uma função chamada `calcula.volume` que calcula o volume da laranja tomando como base o seu diâmetro (considere as laranjas como sendo esféricas) e inclua no data frame `dfLaranja` a informação de volume.

- d) Crie uma função chamada `classifica.destino` que recebe vetores contendo as categorias, pesos e volumes das laranjas e retorna um vetor contendo a classificação de cada uma. Para isso use a seguinte regra para classificação: se a laranja for da categoria A ela deve ser classificada como “descarte”, não devendo ser vendida ou mesmo utilizada para fazer suco; “suco” se a divisão do peso pelo volume for maior que 0,70; “venda”, caso não seja classificada como “descarte” ou “suco”. Acrescente a informação de classificação no data frame `dfLaranja`.
- e) Calcule quantas laranjas pertencem a cada uma das classificações e apresente um gráfico de pizza tomando como base estas informações.
- f) Apresente um histograma de quantidade na cor verde para o índice de acidez das laranjas. O histograma deve ter o título “Histograma do Índice de Acidez” e no eixo X a legenda “Índice de Acidez”.
- g) A partir do histograma do item anterior, faça uma adaptação para que este apresente no eixo Y a densidade de probabilidade e não a quantidade de ocorrências de cada índice de acidez. Neste mesmo histograma, apresente uma função de densidade normal sobreposta.