

# MAC0113 - Introdução à Computação para Ciências Humanas

Prof. Roberto Hirata Jr.

Exercício Programa – Data de entrega: 6/7/2023

## 1 Introdução

A disciplina de Introdução à Programação para Ciências Humanas do curso de Bacharelado em Administração é a primeira disciplina da área de programação do curso. Examinando a página da FEA para a graduação em Administração, encontramos o seguinte texto (os grifos são meus): “O Curso de Administração da FEA-USP é um curso multidisciplinar que qualifica o aluno para **trabalhar em todas as áreas**, com capacidades de **gerenciar** cada uma, bem como, o que é característica única do curso de Administração, **integrar o conhecimento de todas as áreas**. Dessa forma pode-se dizer que o foco do Curso de Administração é **desenvolver a capacidade do aluno para ter uma visão sistêmica na atividade de gestão possibilitando a implementação de técnicas e instrumentos de forma sustentável nos níveis ambiental, social e econômico-financeiro.**”

A disciplina de programação ajuda a alcançar os objetivos acima de várias formas, principalmente no que tange a desenvolver a capacidade do estudante em organizar seu pensamento de forma lógica, exercitando os raciocínios abduutivo, indutivo e dedutivo. Através da programação, o estudante aprende a pensar num problema e a encontrar uma forma de resolvê-lo de forma algorítmica, resolvendo-o por partes, verificando que cada parte está correta, que parte consome mais recursos etc.

Para esta disciplina, para efeitos de avaliação, serão considerados:

- Funcionamento do código. Este item é de fundamental importância para um exercício programa ser considerado entregue. Por funcionamento, entenda-se: **o código apresentado implementa o que foi especificado no enunciado**. Existem várias formas de avaliar este quesito. Uma delas é executar o programa com uma bateria de testes e verificar se o programa passou nos testes, isto é, verificar se os resultados são iguais aos esperados.
- Organização e clareza do código. O código é fácil de ler e entender?
- Documentação do código. As passagens mais difíceis do algoritmo tem frases que ajudem o seu entendimento? As variáveis e constantes estão associadas a frases que dizem para que elas servem?

## 2 O exercício programa (EP)

Especificação matemática: dada uma tabela com o valor do sanduíche Big Mac da rede McDonald's em diversos países e em diversos anos, calcule várias estatísticas e apresente alguns gráficos.

Estimar o valor real de uma moeda é uma tarefa cara e complexa, mas muito útil e lucrativa economicamente. Sabendo esse valor, um investidor pode, por exemplo, fazer melhores compras e vendas nas bolsas de valores internacionais.

O índice Big Mac foi criado pela revista The Economist em 1986 numa “tentativa” didática (“lighthearted”, nas palavras da revista) de ilustrar o valor “correto” de uma moeda. Embora o índice tenha uma base teórica no Poder de Compra Paritário (Purchasing-Power Parity - PPP<sup>1</sup>), essas aspas são justificadas pois existem fatores locais que influenciam o valor de uma mercadoria.

Vamos dar um exemplo de uso do índice Big Mac para o nosso próprio país. O preço médio do Big Mac no Brasil em 1 de julho de 2022 era R\$ 22.9 reais e o valor do Big Mac nos Estados Unidos nesse dia era de US\$ 5.15 dólares. O índice Big Mac é, então, aproximadamente  $\frac{22.9}{5.15} \approx 4.4466$ . Isso significa que o dólar, que valia R\$ 5.39 nesse dia, estaria supervalorizado (ou o real subvalorizado) em  $\frac{4.4466-5.39}{5.39} \approx -0.175$ , ou subvalorizado em 17.5%.

Para quem quiser saber mais sobre este índice, procure no link abaixo.

[https://pt.wikipedia.org/wiki/%C3%8Dndice\\_Big\\_Mac](https://pt.wikipedia.org/wiki/%C3%8Dndice_Big_Mac)

Uma das grandes críticas desse índice é que não leva em conta o valor do trabalho, isto é, o valor pago para o trabalhador da lanchonete. Além disso, há outros fatores que afetam o valor de uma moeda, como a instabilidade política, como conhecemos muito bem no Brasil.

Uma curiosidade, no governo de Cristina Kirchner deu-se tanta importância para este índice que houve pressão para a empresa manter os preços do Big Mac mais baixos (<https://tinyurl.com/24pja3pv>).

Para este EP será fornecido um dos arquivos com os dados disponibilizados pela revista. O nome do arquivo é `ep2Data.csv` e ele contém 1869 linhas (a primeira linha é o cabeçalho) e oito colunas nomeadas (não necessariamente nessa ordem):

- `currencyCode` - uma tripla de letras indicando o código da moeda.
- `date` - data em que o valor do Big Mac foi pesquisado.
- `dollarEx` - o valor do câmbio do dólar no dia.
- `GDPDolar` - Produto Interno Bruto por pessoa (Gross Domestic Product) em dólar.

---

<sup>1</sup>O PPP diz que o valor de troca de uma moeda é proporcional à razão entre os níveis de preço entre dois países.

- GDPLocal - Produto Interno Bruto Doméstico (Gross Domestic Product) em moeda local.
- isoA3 - uma tripla de letras indicando o código do país.
- localPrice - o valor do Big Mac em moeda local.
- name - o nome do país.

As tarefas propostas para este exercício programa são:

**E1.** Faça uma função `criaVetorAno` que recebe um dataframe `D` como parâmetro e retorna um vetor com o ano de cada coleta do valor do Big Mac. Use a seguinte assinatura da função **obrigatoriamente**:

`criaVetorAno < - function(D)`

Dicas:

- dentro da função, comece criando um vetor vazio.
- varra a coluna de data do dataframe e use a função `strsplit` (leia o help da função para saber como ela funciona) para separar a string em três pedaços.
- a função `unlist` transforma uma lista em um vetor e isso facilita sua vida para atribuir a parte do vetor que refere-se ao ano.
- retorne o vetor criado.

**E2.** Faça uma função `BigMacUSPrice` que recebe um dataframe `D` como parâmetro e um número inteiro positivo `A` (que corresponde a um ano) e retorna do valor do Big Mac nos Estados Unidos naquele ano. Use a seguinte assinatura da função **obrigatoriamente**:

`BigMacUSPrice < - function(D,A)`

Dicas:

- varra a coluna de `isoA3` do dataframe e encontre a string `USA` correspondente ao ano `A`. Nesta tabela foram retirados os valores do Big Mac nos Estados Unidos da América que apareciam duas vezes no mesmo ano.
- retorne o valor do Big MAC da linha encontrada.

**E3.** Faça uma função `mediaColunaPais` que recebe como parâmetros um dataframe `D`, um número inteiro positivo `C` (que corresponde a uma coluna **numérica** do dataframe) e uma string de três letras `P` (que corresponde ao código do país) e retorna a média de todos os valores da coluna `C` cujo país é igual a `P`. Você deve usar a fórmula da média sem usar as funções `sum`, ou `mean`, ou qualquer outra que o R tenha pronta. Use a seguinte assinatura da função **obrigatoriamente**:

`mediaColunaPais <- function(D,C,P)`

O valor da média amostral de uma distribuição de  $N$  medidas  $m_i, i \in [1, N]$  é dado por:

$$\hat{m} = \frac{1}{N} \sum_{i=1}^N m_i$$

- E4.** Faça uma função `varColunaPais` que recebe como parâmetros um dataframe `D`, um número inteiro positivo `C` (que corresponde a uma coluna **numérica** do dataframe) e uma string de três letras `P` (que corresponde ao código do país) e retorna a variância de todos os valores da coluna `C` cujo país é igual a `P`. Você deve usar a fórmula da variância sem usar as funções `sum`, ou `mean`, ou `var`, ou qualquer outra que o R tenha pronta. Use a seguinte assinatura da função **obrigatoriamente**:

`varColunaPais <- function(D,C,P)`

O valor da variância amostral de uma distribuição de  $N$  medidas  $m_i, i \in [1, N]$  é dado por:

$$\hat{s}^2 = \frac{1}{N-1} \sum_{i=1}^N (m_i - \hat{m})^2$$

- E5.** Faça uma função `criaVetorBMI` que recebe um dataframe `D` como parâmetro e retorna um vetor com o índice Big Mac de cada coleta do valor do Big Mac. Use a seguinte assinatura da função **obrigatoriamente**:

`criaVetorBMI <- function(D)`

Dicas:

- dentro da função, comece criando um vetor vazio.
- chama a função `BigMacUSPrice` para cada ano de 2000 a 2022 e coloca nesse primeiro vetor.
- crie um novo vetor vazio
- varra a coluna `localPrice` do dataframe e faça a conta do índice, levando em conta o preço do Big Mac nos Estados Unidos naquele ano escolhido.
- retorne o vetor criado.

- E6.** Faça um programa em R que:

- Leia o arquivo de dados de nome `epData.csv`. Para isso, veja a ajuda (“help”) do comando `read.csv` e preste atenção no uso do separador (`sep`). Neste caso o arquivo fornecido é `comma-separated-values (csv)`, mas nem sempre é assim. Algumas vezes o arquivo tem o final `csv` mas é `semicolon-separated-values`. Nesses casos a leitura deve ser atribuída a uma variável chamada `BigMacData`. Por exemplo:

```
BigMacData = read.csv(<nome do arquivo>, sep=<separador>)
```

Use o comando `ls.strBigMacData` para ver se o arquivo foi lido corretamente. O comando retorna o rótulo de cada coluna e alguns dados daquela coluna, como abaixo.

```
> ls.str(BigMacData)
currencyCode : chr [1:1946] "ARS" "AUD" "BRL" "CAD" "CHF" ...
date : chr [1:1946] "2000-04-01" "2000-04-01" "2000-04-01" ...
dollarEx : num [1:1946] 1 1.68 1.79 1.47 1.7 ...
GDPDollar : num [1:1946] 8709 21747 3501 22341 41859 ...
GDPLocal : num [1:1946] 8709 33699 6351 33191 62879 ...
isoA3 : chr [1:1946] "ARG" "AUS" "BRA" "CAN" "CHE" "CHL" ...
localPrice : num [1:1946] 2.5 2.59 2.95 2.85 5.9 ...
name : chr [1:1946] "Argentina" "Australia" "Brazil" ...
```

- Use a função `criaVetorAno` e agregue o resultado ao dataframe `BigMacData` usando a função `cbind`.
- Use a função `criaVetorBMI` e agregue o resultado ao dataframe `BigMacData` usando a função `cbind`.

**Exercício intermediário.** Use o código abaixo para facilitar ter os códigos de todos os países que aparecem no arquivo:

```
> Países <- unique(sort(BigMacData[,"isoA3"]))
> Países
 [1] "ARE" "ARG" "AUS" "AUT" "AZE" "BEL" "BHR" "BRA" "CAN" "CHE"
[11] "CHL" "CHN" "COL" "CRI" "CZE" "DEU" "DNK" "EGY" "ESP" "EST"
[21] "FIN" "FRA" "GBR" "GRC" "GTM" "HKG" "HND" "HRV" "HUN" "IDN"
[31] "IND" "IRL" "ISR" "ITA" "JOR" "JPN" "KOR" "KWT" "LBN" "LKA"
[41] "LTU" "LVA" "MDA" "MEX" "MYS" "NIC" "NLD" "NOR" "NZL" "OMN"
[51] "PAK" "PER" "PHL" "POL" "PRT" "QAT" "ROU" "RUS" "SAU" "SGP"
[61] "SVK" "SVN" "SWE" "THA" "TUR" "TWN" "UKR" "URY" "USA" "VNM"
[71] "ZAF"
```

Crie dois outros vetores de agrupamentos de países de acordo com algum critério decidido por você. No relatório, justifique a decisão. Por exemplo, países agrupados por continente. Quanto mais criativa e bem justificada a escolha, melhor. Esses dois vetores serão usados para fazer gráficos por países.

- E7.** Faça um gráfico do índice Big Mac médio dos países (eixo *Y*). No eixo *X* os países de acordo com a variável `Países`. Faça mais dois gráficos do índice Big Mac médio usando os vetores criados no item anterior no eixo *X*. Explique o que você vê nos gráficos no seu relatório.

**E8.** Faça um gráfico da variância do índice Big Mac dos países (eixo  $Y$ ). No eixo  $X$  os países de acordo com a variável `Países`. Faça mais dois gráficos da variância do índice Big Mac usando os vetores criados no item anterior no eixo  $X$ . Explique o que você vê nos gráficos no seu relatório.

Há duas entregas do EP:

- A primeira entrega consistirá no envio (“upload”), via e-disciplinas, de dois arquivos, ou três arquivos (se você escolheu fazer o relatório em RMarkdown). Um deles é o código em R para resolver o problema e o outro arquivo é o relatório em PDF com os gráficos. Se você fez o relatório em RMarkdown, você terá que enviar um arquivo `.Rmd`.

Não esqueça de colocar seu nome e número USP no EP e no relatório, pois trabalhos sem essas informações não serão corrigidos.

- A segunda entrega consistirá no envio (“upload”), via e-disciplinas, do arquivo `.R` no VPL (sistema para avaliação automática de parte do EP)

### 3 Plágio

Plágio é a cópia/modificação não autorizada e/ou sem o conhecimento do autor original (inclusive de um autor não humano, como o chatGPT). O plágio é um problema grave que pode levar até a expulsão do estudante da universidade.

### 4 Observações:

1. A endentação correta do programa será considerada.
2. O exercício é **individual**. Exercícios que indiquem o contrário serão tratados como quebra do código de ética discente da USP, receberão nota 0 (zero) e o caso será levado ao conhecimento da Câmara de Graduação da FEA.
3. O cabeçalho do programa deve ser **obrigatoriamente** como abaixo. Não esqueça de colocar o seu nome completo, o seu número USP, e a sua turma (01, 02, 21 ou 22).

```
#####  
## AO PREENCHER ESSE CABEÇALHO COM O MEU NOME E O MEU NÚMERO USP,  
## DECLARO QUE SOU O ÚNICO AUTOR E RESPONSÁVEL POR ESSE PROGRAMA.  
## TODAS AS PARTES ORIGINAIS DESSE EXERCÍCIO PROGRAMA (EP) FORAM  
## DESENVOLVIDAS E IMPLEMENTADAS POR MIM SEGUINDO AS INSTRUÇÕES
```

```
## DESSE EP E QUE PORTANTO NÃO CONSTITUEM DESONESTIDADE ACADÊMICA
## OU PLÁGIO.
## DECLARO TAMBÉM QUE SOU RESPONSÁVEL POR TODAS AS CÓPIAS
## DESSE PROGRAMA E QUE EU NÃO DISTRIBUI OU FACILITEI A
## SUA DISTRIBUIÇÃO. ESTOU CIENTE QUE OS CASOS DE PLÁGIO E
## DESONESTIDADE ACADÊMICA SERÃO TRATADOS SEGUNDO OS CRITÉRIOS
## DIVULGADOS NA PÁGINA DA DISCIPLINA.
## ENTENDO QUE EPS SEM ESTE CABEÇALHO NÃO SERÃO CORRIGIDOS E,
## AINDA ASSIM, PODERÃO SER PUNIDOS POR DESONESTIDADE ACADÊMICA.
```

```
## Nome :
## NUSP :
## Turma:
## Prof.: Roberto Hirata Jr.
```

```
## Referências: Com exceção das rotinas fornecidas no enunciado
## e em sala de aula, caso você tenha utilizado alguma referência,
## liste-as abaixo para que o seu programa não seja considerado
## plágio ou irregular.
```

```
## Exemplo:
## - O algoritmo Quicksort foi baseado em
## http://wiki.python.org.br/QuickSort
#####
```

```
## [Seu programa]
```