

Relatório do EP - MAC0113

Jonatas Vieira NUSP: 9913690 T21

2023-07-06

Introdução

O Índice Big Mac é um indicador econômico não convencional que compara o poder de compra de diferentes moedas. Ele usa o preço de um hambúrguer Big Mac, um produto globalmente padronizado, para revelar as disparidades das taxas de câmbio e avaliar as avaliações da moeda. Embora tenha limitações, o Índice Big Mac fornece informações sobre a paridade do poder de compra e é amplamente discutido na economia internacional. Este projeto visa criar um código, em R, onde seja possível estimar o índice Big Mac em diferentes anos e países, bem como definir gráficos comparativo desses índices para cada país. Além disso, a situação problema também tem como objetivo colocar em prática os assuntos abordados em aula, na disciplina MAC0113 - Introdução à Computação para Ciências Humanas, colocando em prática o uso do conhecimento em um problema real.

Dados

Os dados que serviram de base para o problema foram concedidos pelo professor da disciplina, que por sua vez fez a extração dos dados disponibilizados pela revista The Economist, e refletem o valor do sanduíche Big Mac da rede McDonald's em diversos países e em diversos anos. Além disso, trazem informações como valor do câmbio do dólar no dia e o Produto Interno Bruto por pessoa (Gross Domestic Product) em dólar, conforme mostra a Tabela 1, que contém uma parte dos dados da base.

Table 1: Tabela 1

name	isoA3	currencyCode	localPrice	dollarEx	GDPDollar	GDPLocal	date
Argentina	ARG	ARS	2.50	1.00	8709.072	8709.072	2000-04-01
Australia	AUS	AUD	2.59	1.68	21746.809	33698.764	2000-04-01
Brazil	BRA	BRL	2.95	1.79	3501.438	6351.375	2000-04-01
Canada	CAN	CAD	2.85	1.47	22340.553	33191.468	2000-04-01

Os dados estão distribuídos em um em um arquivo no formato .csv com 8 colunas, sendo que, para cada uma delas, é possível identificar o formato do dado associado, o número de observações e a quantidade de variáveis, por exemplo, conforme a Figura 1.

Figura 1:

```
## 'data.frame': 1868 obs. of 8 variables:
## $ name : chr "Argentina" "Australia" "Brazil" "Canada" ...
## $ isoA3 : chr "ARG" "AUS" "BRA" "CAN" ...
## $ currencyCode: chr "ARS" "AUD" "BRL" "CAD" ...
## $ localPrice : num 2.5 2.59 2.95 2.85 5.9 ...
## $ dollarEx : num 1 1.68 1.79 1.47 1.7 ...
```

```
## $ GDPDollar : num 8709 21747 3501 22341 41859 ...
## $ GDPLocal : num 8709 33699 6351 33191 62879 ...
## $ date : chr "2000-04-01" "2000-04-01" "2000-04-01" "2000-04-01" ...
```

Método

Para resolver o problema proposto, foram criadas algumas funções, que recebem parâmetros de entrada, executam operações e retornam valores de resultado, conforme mostrado abaixo. Com isso, as funções construídas foram utilizadas em vários trechos do EP, sem a necessidade de codificá-la novamente.

A função da Figura 2, `criaVetorAno`, foi criada para receber um dataframe `D` como parâmetro e retornar um vetor com o ano de cada coleta do valor do Big Mac. Nela a data no formato DD-MM-AAAA, presente a coluna “date” da base de dados, foi dividida em “DD”, “MM” e “AAAA”, com o auxílio da função `strsplit` e, posteriormente guardada em um vetor, com o auxílio da função `unlist`. Por fim, apenas o ano “AAAA” foi guardado em um novo vetor, retornando esse como resultado da função.

Figura 2:

```
criaVetorAno <- function(D){
  dataSeparada <- character()
  vetorAno <- character()
  for(i in 1:nrow(D)){
    dataSplit <- strsplit(D[i,"date"],split = "-")
    dataSeparada <- unlist(list(dataSeparada,dataSplit))
  }
  #Colocando apenas o ano (AAAA) no vetorAno
  for (i in seq(from = 1, to = length(dataSeparada), by = 3)) {
    vetorAno <- c(vetorAno, dataSeparada[i])
  }
  return(vetorAno)
}
```

A função da Figura 3, `BigMacUSPrice`, foi criada para receber um dataframe `D` e um ano `A` como parâmetro e retornar valor do Big Mac nos Estados Unidos naquele ano. Nela, foi analisada a coluna com o dado sobre a tripla de letras indicando o código do país, “isoA3”, e, para os casos em que o código era “USA”, a coluna “date” foi analisada com o auxílio da função da Figura 2, `criaVetorAno`, comparando o ano recebido como parâmetro do dataframe com o ano gerado pela função da Figura 2 e, caso esses anos fossem iguais, o valor da coluna “localPrice” é retornado como resultado da função.

Figura 3:

```
BigMacUSPrice <- function(D,A){
  for(i in 1:nrow(D)){
    if(D[i,"isoA3"]=="USA"){
      vetorAno <- criaVetorAno(D)
      if(vetorAno[i] == A){
        precoBigMacEUA <- D[i,"localPrice"]
      }
    }
  }
  return(precoBigMacEUA)
}
```

A função da Figura 4, `mediaColunaPais`, foi criada para receber um dataframe `D`, um código de país `P` e um número correspondente a uma coluna do dataframe `C` como parâmetro e retorna a média de todos os

valores da coluna C cujo país é igual a P. Ela se inicia com a verificação da classe da coluna C e, caso fosse numérica, iniciaria o cálculo da média amostral, para as ocorrências em que a coluna “isoA3” fosse igual ao código do país de entrada como parâmetro da função. Esse cálculo foi realizado com o auxílio de um laço, que somava os valores da coluna C e, após percorrer as linhas do dataframe, dividia os valores somados pelo número de ocorrências. Caso a coluna C fosse numérica, o valor da média para essa de um dado país P é retornado como resultado da função. Caso contrário, o valor retornado é nulo.

Figura 4:

```
mediaColunaPais <- function(D,C,P){
  #Verificando se a classe da coluna C é numérica
  if(class(D[, C]) == "numeric"){
    soma <- as.numeric(0)
    nAmostras <- as.numeric(0)
    for(i in 1:nrow(D)){
      if(D[i,"isoA3"] == P){
        soma <- soma + D[i,C]
        nAmostras <- nAmostras + 1
      }
    }
    media <- soma/nAmostras
  }else{
    media <- NULL
  }
  return(media)
}
```

A função da Figura 5, varColunaPais, foi criada para receber um dataframe D, um código de país P e um número correspondente a uma coluna do dataframe C como parâmetros e retorna a média de todos os valores da coluna C cujo país é igual a P. Ela se inicia com a verificação da classe da coluna C e, caso fosse numérica, iniciaria o cálculo da média amostral, com o auxílio da função da Figura 4. Para calcular a variância, para as ocorrências em que a coluna “isoA3” fosse igual ao código do país de entrada como parâmetro da função, foi utilizado um laço, que somava os valores do quadrado da subtração entre os valores coluna C e a média gerada pela função da figura 4 e, após percorrer as linhas do dataframe, dividia os valores somados pelo numero de ocorrencias - 1. Caso a coluna C fosse numérica, o valor da variância para essa de um dado país P é retornado como resultado da função. Caso contrário, o valor retornado é nulo.

Figura 5:

```
varColunaPais <- function(D,C,P){
  if(class(D[, C]) == "numeric"){
    soma <- as.numeric(0)
    nAmostras <- as.numeric(0)
    mediaAmostral <- mediaColunaPais(D,C,P)
    for(i in 1:nrow(D)){
      if(D[i,"isoA3"] == P){
        soma <- soma + (D[i,C] - mediaAmostral)^2
        nAmostras <- nAmostras + 1
      }
    }
    variancia <- soma/(nAmostras-1)
  }else{
    variancia <- NULL
  }
}
```

```

    return(variancia)
}

```

A função da Figura 6, `criaVetorBMI`, foi criada para receber um dataframe `D` como parâmetro e retornar um vetor com o índice Big Mac de cada coleta do valor do Big Mac. Ela se inicia com a criação de um vetor, por meio do auxílio de um laço e da função da Figura 3, com todos os preços do Big Mac nos Estados Unidos entre os anos de 2000 e 2022. Após isso, um novo vetor foi criado, `anoPreco`, por meio do uso da função da Figura 2 e, com isso, foi possível criar um laço com o número de passagens igual ao número de linhas do dataframe. Também foi utilizada uma variável auxiliar, que tinha como objetivo garantir que todos os valores acessados em cada passo do laço fossem correspondentes a mesma linha do dataframe. Uma comparação entre o `anoPreco` foi realizada com a variável auxiliar `anoAnalisado`, que parte do ano 2000, a fim de garantir que o valor da coluna “localPrice” seria correspondente ao ano de uma dada linha do dataframe e, caso a comparação seja verdadeira, o Índice Big Mac era calculado por meio da divisão do “localPrice” da linha do dataframe e do preço do Big Mac nos Estados Unidos, no mesmo ano correspondente ao “localPrice” acessado. Vale ressaltar que para garantir que essa divisão fosse correta, foi criada uma variável auxiliar `idPrecoEUA`, que tem como função acessar o valor exato do preço do Big Mac nos Estados Unidos para o ano correspondente ao preço local do Big Mac em um dado país. Caso a comparação entre `anoPreco` e `anoAnalisado` fosse falsa, o preço do Big Mac acessado era o do próximo ano, garantindo o cálculo do índice correto. Por fim, cada índice calculado era gravado em um vetor, que seria retornado como resultado da função.

Figura 6:

```

criaVetorBMI <- function(D){
  precoEUA <- numeric()
  #acessando cada valor do Big Mac para um ano entre 2000 e 2022
  for(ano in 2000:2022){
    precoEUA <- c(precoEUA, BigMacUSPrice(D, ano))
  }
  anoAnalisado <- 2000
  idPrecoEUA <- 1
  contador <- 1
  anoPreco <- criaVetorAno(D)
  indiceBigMac <- numeric()

  for(i in 1:length(anoPreco)){
    #compara se o ano de um dado é o mesmo do analisado
    if(anoPreco[contador] == (anoAnalisado) && anoAnalisado <= "2022"){
      indice <- D[contador,"localPrice"]/precoEUA[idPrecoEUA]
      indiceBigMac <- c(indiceBigMac, indice)
      contador <- contador+1}else{
      #caso o ano do dado for diferente, calcula levando em conta o valor
      #do Big Mac do proximo ano
      anoAnalisado <- anoAnalisado+1
      idPrecoEUA <-idPrecoEUA+1
      indice <- D[contador,"localPrice"]/precoEUA[idPrecoEUA]
      indiceBigMac <- c(indiceBigMac, indice)
      contador <- contador+1
    }
  }
  return(indiceBigMac)
}

```

A Figura 7 mostra outras duas funções criadas com intuito de reaproveitar o código, ou seja, nela é possível

criar um vetor de variância (`varianciaPaíses`) e índice Big Mac médio dos países (`indiceBigMacMedioPais`) a partir dos parâmetros de entrada `países`, que é um vetor com os códigos dos países, `coluna`, que representa a coluna onde será feita a análise estatística, e o `dataframe`. A ideia surgiu com a necessidade de fazer esse cálculos 3 vezes para construir os gráficos pedidos no problema, ou seja, utilizar 3 vezes para cada uma das estatísticas a mesma estrutura de código. Porém, com o uso das funções, a estrutura foi codificada uma única vez. Os vetores de variância e índice Big Mac médio são retornados como resultado dessas funções para uma dada lista de código de países.

Figura 7:

```
#Reaproveitar o código do calculo da média do IBM para um vetor de países
indiceBigMacMedioPais <- function(países,coluna,dataframe){
  indiceBigMacMedio <- numeric()
  for(i in 1:length(países)){
    indiceBigMacMedio <- c(indiceBigMacMedio, mediaColunaPais(dataframe,coluna,países[i]))
    indiceBigMacMedio <-round(indiceBigMacMedio, 3)
  }
  return(indiceBigMacMedio)
}

#Reaproveitar o código do calculo da variância do IBM para um vetor de países
varianciaPaíses <- function(países,coluna,dataframe){
  varPaíses <- numeric()
  for(i in 1:length(países)){
    varPaíses <- c(varPaíses, varColunaPais(dataframe,coluna,países[i]))
    varPaíses <-round(varPaíses, 3)
  }
  return(varPaíses)
}
```

Para iniciar o programa, foi solicitado ao usuário o caminho de origem do arquivo .csv, que foi lido conforme mostra a Figura 8.

Figura 8:

```
cat("Digite o caminho da base se dados: ")
caminhoCSV <- readline()
#past concatena a entrada do usuário com as aspas para leitura do arquivo
BigMacData <- read.csv(paste(caminhoCSV, sep = ""), header = TRUE)
```

Após a leitura do arquivo, foram adicionadas duas colunas ao `dataframe`, com o uso da função `cbind`, contendo informações de ano e índice Big Mac, geradas pelo uso das funções das Figuras 2 e 6, respectivamente, conforme a Figura 9.

Figura 9:

```
#Adiciona 2 colunas de dados ao dataframe
BigMacData <- cbind(BigMacData, Ano = criaVetorAno(BigMacData))
BigMacData <- cbind(BigMacData, indiceBM = criaVetorBMI(BigMacData))
```

Por fim, a Figura 10 mostra a estrutura utilizada para gerar os gráficos solicitados no problema, com o uso das funções da Figura 7.

Figura 10:

```

#Gráficos para todos os países do dataframe
barplot(indiceBigMacMedioPais(Paises,10,BigMacData),xlab= "Países",ylab="IBM médio",
        main="IBM médio dos Países", names.arg = Paises, cex.names = 0.3, las =2)

barplot(varianciaPaises(Paises,10,BigMacData),xlab= "Países",ylab="Variância",
        main="Variância do IBM dos Países", names.arg = Paises, cex.names = 0.3, las =2)

#Gráfico para os países que são classificados como "economias avançadas"(segundo o FMI)
barplot(indiceBigMacMedioPais(paisesEconomiaAvancada,10,BigMacData),xlab= "Países",ylab="IBM médio",
        main="IBM médio dos Países de economia avançada",
        names.arg = paisesEconomiaAvancada, cex.names = 0.3, las =2)

barplot(varianciaPaises(paisesEconomiaAvancada,10,BigMacData),xlab= "Países",ylab="Variância",
        main="Variância do IBM dos Países de economia avançada",
        names.arg = paisesEconomiaAvancada, cex.names = 0.3, las =2)

#Gráfico para os outros países do dataframe
barplot(indiceBigMacMedioPais(outrosPaises,10,BigMacData),xlab= "Países",ylab="IBM médio",
        main="IBM médio dos Países de economia não avançada",
        names.arg = outrosPaises, cex.names = 0.3, las =2)

barplot(varianciaPaises(outrosPaises,10,BigMacData),xlab= "Países",ylab="Variância",
        main="Variância do IBM dos Países de economia não avançada",
        names.arg = outrosPaises, cex.names = 0.3, las =2)

```

Resultados

A Figura 11 mostra que o arquivo foi lido corretamente.

Figura 11:

```

## currencyCode : chr [1:1868] "ARS" "AUD" "BRL" "CAD" "CHF" "CLP" "CNY" "CZK" "DKK" "GBP" ...
## date : chr [1:1868] "2000-04-01" "2000-04-01" "2000-04-01" "2000-04-01" ...
## dollarEx : num [1:1868] 1 1.68 1.79 1.47 1.7 ...
## GDPDollar : num [1:1868] 8709 21747 3501 22341 41859 ...
## GDPLocal : num [1:1868] 8709 33699 6351 33191 62879 ...
## isoA3 : chr [1:1868] "ARG" "AUS" "BRA" "CAN" "CHE" "CHL" "CHN" "CZE" "DNK" "GBR" ...
## localPrice : num [1:1868] 2.5 2.59 2.95 2.85 5.9 ...
## name : chr [1:1868] "Argentina" "Australia" "Brazil" "Canada" "Switzerland" ...

```

A Figura 12 mostra os agrupamentos para os vetores Paises, paisesEconomiaAvancada e outrosPaises, sendo o primeiro criado de acordo com as instruções do problema e o segundo e terceiro criado com base em dados do Fundo Monetário Internacional (FMI).

Figura 12:

```

## vetor Paises:

## [1] "ARE" "ARG" "AUS" "AUT" "AZE" "BEL" "BHR" "BRA" "CAN" "CHE" "CHL" "CHN"
## [13] "COL" "CRI" "CZE" "DEU" "DNK" "EGY" "ESP" "EST" "FIN" "FRA" "GBR" "GRC"
## [25] "GTM" "HKG" "HND" "HRV" "HUN" "IDN" "IND" "IRL" "ISR" "ITA" "JOR" "JPN"
## [37] "KOR" "KWT" "LBN" "LKA" "LTU" "LVA" "MDA" "MEX" "MYS" "NIC" "NLD" "NOR"
## [49] "NZL" "OMN" "PAK" "PER" "PHL" "POL" "PRT" "QAT" "ROU" "RUS" "SAU" "SGP"
## [61] "SVK" "SVN" "SWE" "THA" "TUR" "TWN" "UKR" "URY" "USA" "VNM" "ZAF"

```

```
## vetor paiseseconomiaAvancada:
```

```
## [1] "AUS" "AUT" "BEL" "CAN" "KOR" "DNK" "SVK" "SVN" "ESP" "USA" "EST" "FIN"
## [13] "FRA" "GRC" "HKG" "IRL" "ISR" "ITA" "JPN" "LVA" "LTU" "NZL" "NOR" "NLD"
## [25] "PRT" "GBR" "CZE" "SGP" "SWE" "CHE" "TWN"
```

```
## vetor outrosPaises:
```

```
## [1] "ARE" "ARG" "AZE" "BHR" "BRA" "CHL" "CHN" "COL" "CRI" "DEU" "EGY" "GTM"
## [13] "HND" "HRV" "HUN" "IDN" "IND" "JOR" "KWT" "LBN" "LKA" "MDA" "MEX" "MYS"
## [25] "NIC" "OMN" "PAK" "PER" "PHL" "POL" "QAT" "ROU" "RUS" "SAU" "THA" "TUR"
## [37] "UKR" "URY" "VNM" "ZAF"
```

Para os países apresentados no dataframe, foi gerado o gráfico da Figura 13, mostrando o IBM médio dos Países. É possível observar que há pontos extremos, como o da Turquia, o que é confirmado pela variância do IBM dos Países, como mostra o gráfico da Figura 14.

Figura 13:

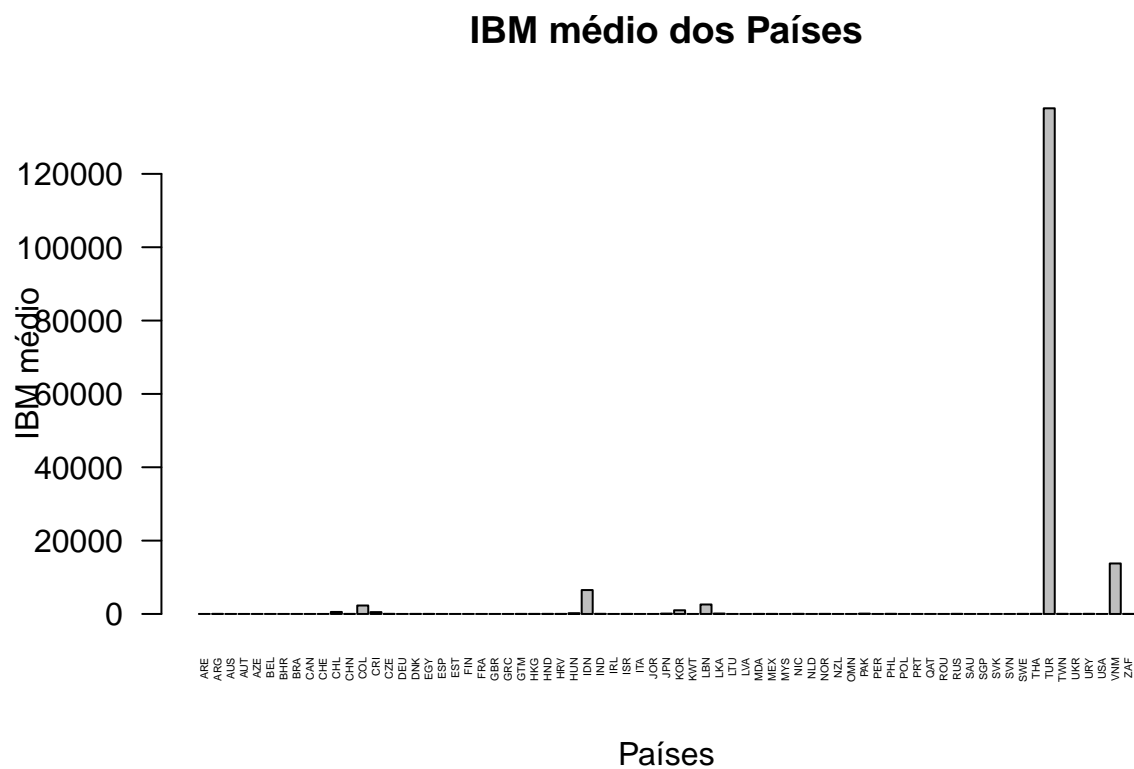
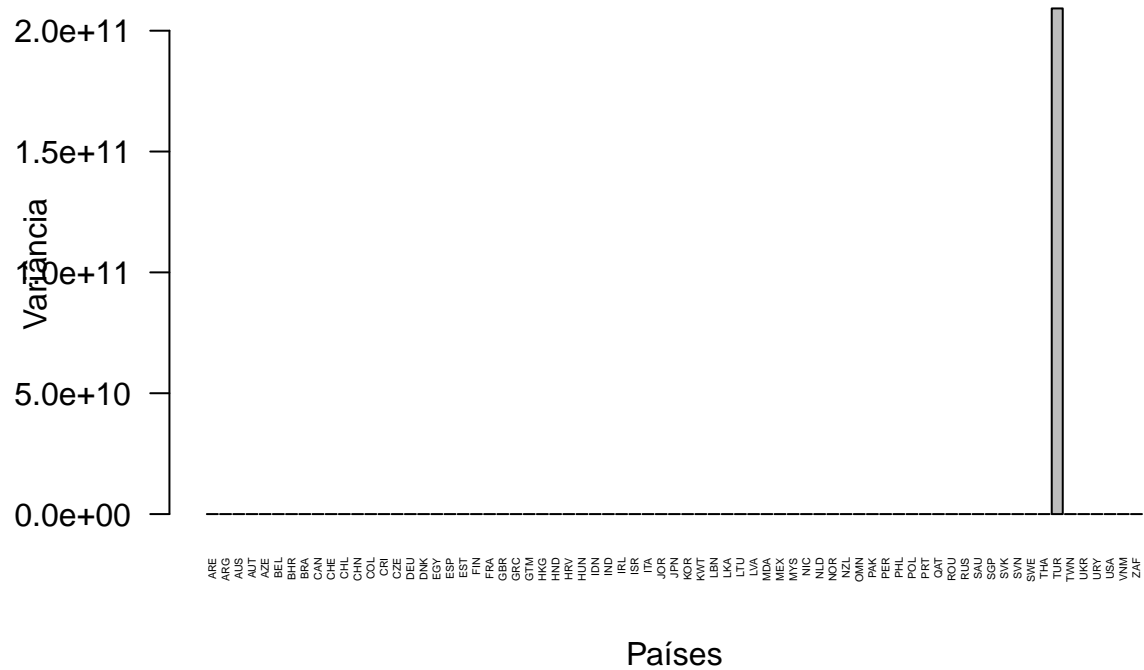


Figura 14:

Variância do IBM dos Países



Para os países de economia avançada, segundo o FMI, foi gerado o gráfico da Figura 15, mostrando o IBM médio. É possível observar que há pontos extremos, mesmo em economias avançadas, como o da Coreia do Sul, Japão e Taiwan. Ao analisar a variância do IBM países de economia avançada, é possível ver que o Japão e Coreia do Sul são pontos fora da curva, como mostra o gráfico da Figura 16, o que evidencia o distanciamento do valor médio do IBM para esses países.

Figura 15:

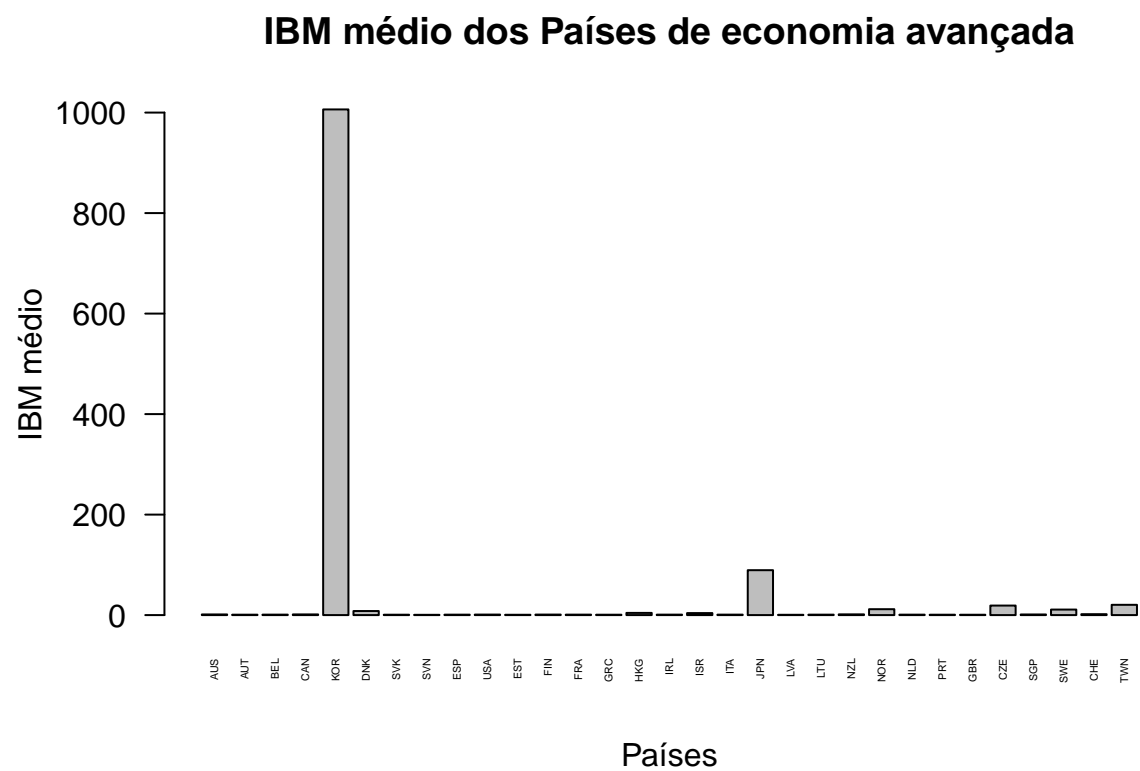
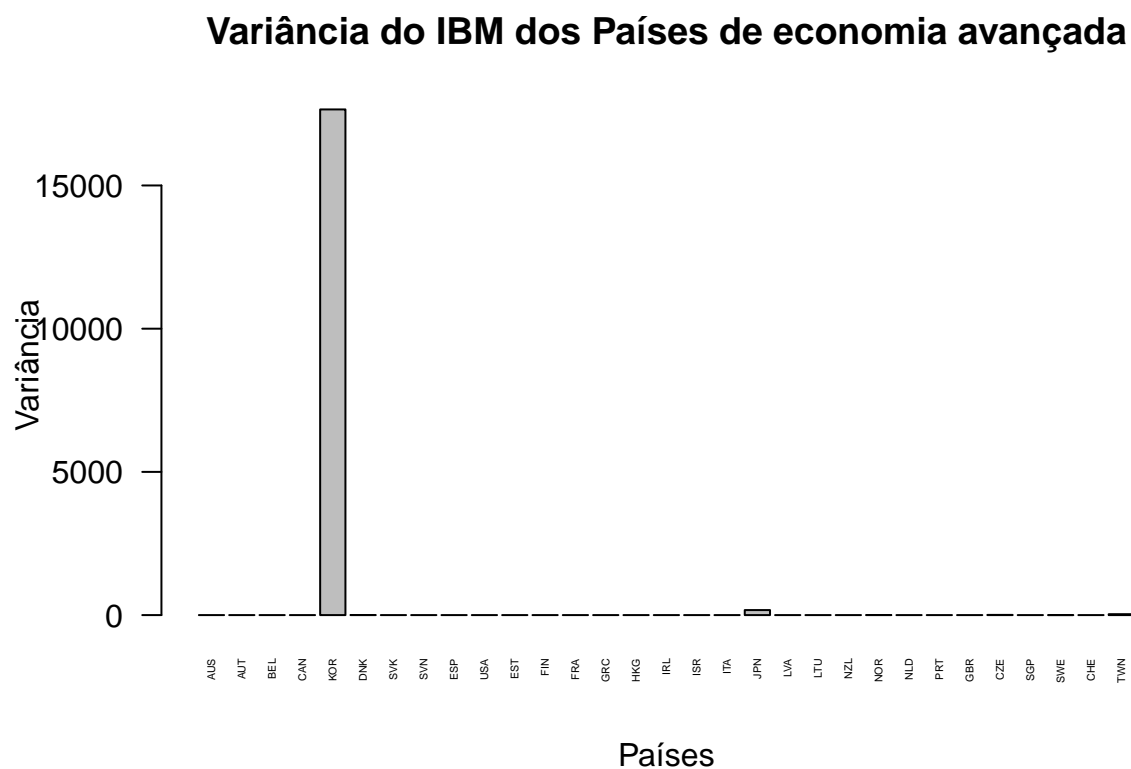


Figura 16:



Para os países de economia não avançada, segundo o FMI, foi gerado o gráfico da Figura 17, mostrando o IBM médio. É possível observar que há pontos extremos, mais destacados que os valores do gráfico equivalente anterior, como Turquia, Vietnã, Índia e Colômbia. Porém, ao analisar a variância do IBM países de economia não avançada, é possível ver que o destaque fica para Turquia, como mostra o gráfico da Figura 18, o que evidencia o distanciamento significativo dos valores de IBM, ao longo dos anos, do valor médio do IBM para esse país.

Figura 17:

IBM médio dos Países de economia não avançada

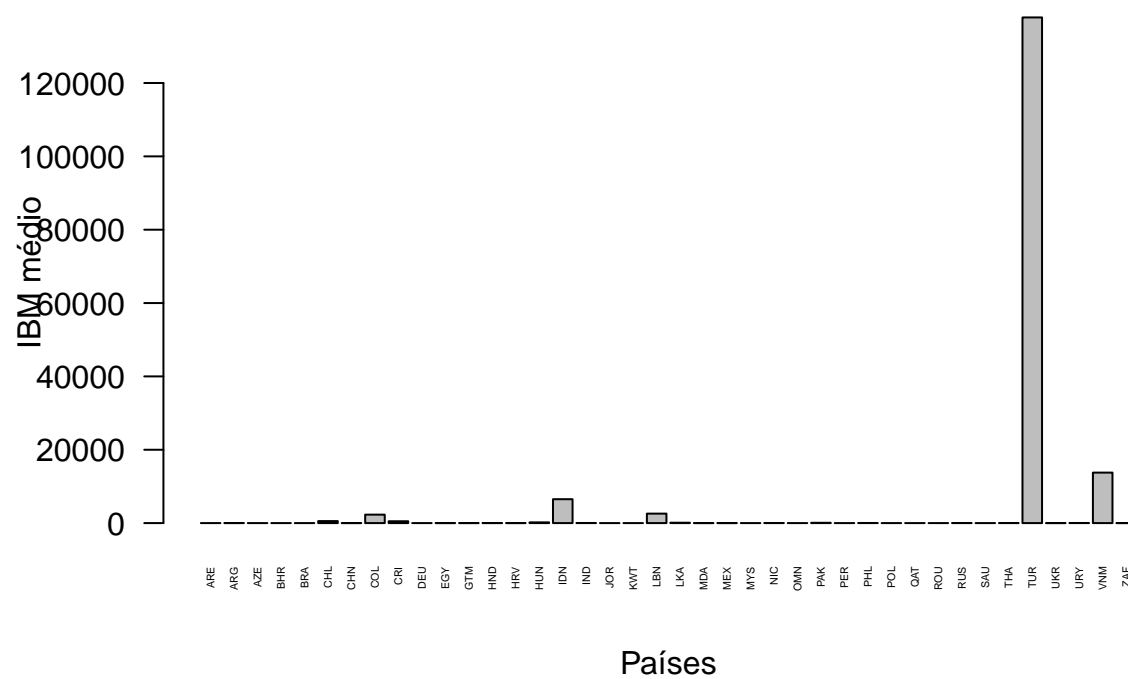
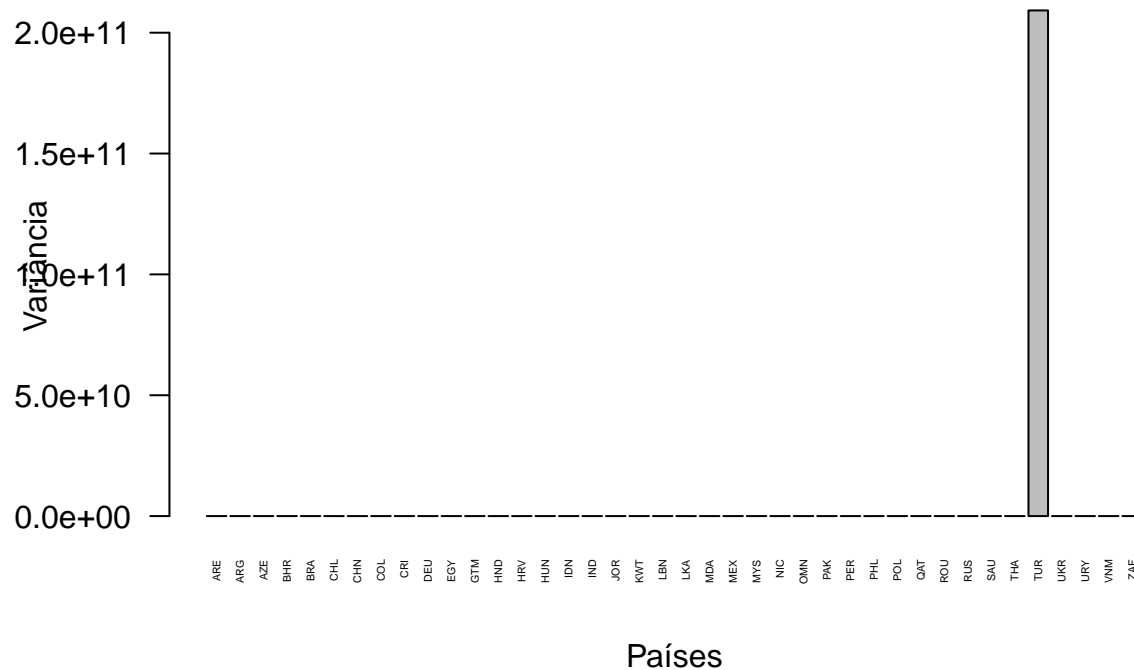


Figura 18:

Variância do IBM dos Países de economia não avançada



Conclusão

É possível concluir que o índice Big Mac (IBM) é um indicador que ilustra bem a disparidade do câmbio do dólar para moedas de outros países. Alguns casos ficaram em destaque, conforme mostrado nos resultados, principalmente para países de economia não avançada, como a Turquia, Vietnã, Índia e Colômbia. Porém, vale destacar que para os países de economia avançada, há também países com alta disparidade de câmbio em relação ao dólar, como Coreia do Sul, Japão e Taiwan. A variância do IBM, principalmente para a Coreia do Sul e Japão, reforçam ainda mais a volatilidade da moeda desses países, além do baixo poder de compra no mercado internacional.

O a situação problema possibilitou, também, a consolidação das dinâmicas aprendidas em aula, na medida em que colocou em prática os seguintes pontos da disciplina:

- Uso de conceitos looping no R;
- Criação de condições no R;
- Conceito de funções e reuso de código;
- Como transferir um problema real para linguagem computacional;

Esses e outros conceitos técnicos foram fundamentais para obter o resultado esperado pelo exercício.

Por fim, para uma próxima versão do código, faz-se necessário a avaliação do IBM com a retirada dos países geradores de pontos extremos nos gráficos, além de uma reordenação para que seja levado em consideração os países com moedas que estão valorizadas em relação ao dólar no período de análise, possibilitando, assim, uma avaliação mais completa do IBM para todos os países que compõe o dataframe.