



SÉRIE DE NOTAS TÉCNICAS EM ECONOMIA DA UFG

NT N. 14

NOTA TÉCNICA EM ECONOMIA **n. 14**

Publicação cujo objetivo é auxiliar na elaboração de aulas e de pesquisas do Curso de Ciências Econômicas da UFG, divulgando demonstrações técnicas e metodológicas ou rotinas computacionais voltadas para a economia.

As opiniões contidas nesta publicação não representam o ponto de vista do Curso ou da FACE/UFG, sendo de inteira responsabilidade de seu(s) autor(es).

Reprodução permitida, desde que citada a fonte.

FACE – Faculdade de Administração, Ciências Contábeis e Ciências Econômicas
Curso de Ciências Econômicas
Direção FACE

Profª. Andrea Freire de Lucena

Vice-Direção FACE

Profª. Daiana Paula Pimenta

Coordenação do Curso de Ciências Econômicas

Prof. Everton Sotto Tibiriçá Rosa

Extração de dados do Sistema Gerenciador de Séries Temporais do Bacen em linguagem R

Antonio Leonardo de Azevedo Carvalho
PPGECON/FACE/UFG

Raphael Maciel de Lima
PPGECON/FACE/UFG

Tatiane Ferreira do Nascimento Melo
PPGECON e IME/UFG

FACE/UFG
Goiânia – Dezembro de 2023
Versão 1.0

Endereço

Campus Samambaia, Prédio da FACE – Rodovia
Goiânia/Nova Veneza, km. 0, CEP 74690-900,
Goiânia – GO. Tel. (62) 3521 – 1390

URL

<http://www.eco.face.ufg.br>

**CIÊNCIAS
ECONÔMICAS**

FACE
FACULDADE DE
ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS E
CIÊNCIAS ECONÔMICAS



Dados Internacionais de Catalogação na Publicação (CIP)
GPT/BC/UFG

Carvalho, Antonio Leonardo de Azevedo. Extração de dados do Sistema Gerenciador de Séries Temporais do Bacen em linguagem R / Lima, Raphael Maciel de. Melo, Tatiane Ferreira do Nascimento. - 2023.
13 f.: (Série de Notas Técnicas em Economia, 014)
Universidade Federal de Goiás, Goiânia, 2023. Bibliografia.
Anexos. 1. Software R. 2. Bacen. 3. Dados. I. Título.

Extração de dados do Sistema Gerenciador de Séries Temporais do Bacen em linguagem R

Antonio Leonardo de Azevedo Carvalho¹

PPGECON/FACE/UFG

Raphael Maciel de Lima²

PPGECON/FACE/UFG

Tatiane Ferreira do Nascimento Melo³

PPGECON e IME/UFG

Resumo: Esta nota técnica tem como objetivo apresentar as principais ferramentas para importar dados do SGS – Sistema Gerenciador de Séries Temporais do Banco Central do Brasil por meio da linguagem de programação *R* usando o pacote *rbcb*.

Palavras chave: R, ggplot, rbc, IPCA, PIB, Câmbio

Abstract: *This technical note aims to present the main tools for importing data from the SGS - Time Series Manager System of the Central Bank of Brazil using the R programming language with the rbc package.*

Key words: R, rbc, ggplot, IPCA, GDP, exchange

¹ antonio.leonardo@discente.ufg.br

² raphael.lima@discente.ufg.br

³ tmelo@ufg.br

1. Introdução

Um dos principais problemas para o pesquisador é a escolha de um banco de dados e o conhecimento de como importar os dados via Internet. Apesar de ser fácil o acesso a alguns dados econômicos, outros necessitam de uma expertise para obtê-los e processá-los de forma eficiente.

A presente nota técnica tem a finalidade de explicar, por meio da linguagem de programação *R* (R Core Team, 2023), uma maneira de baixar dados econômicos com maior facilidade. Utiliza-se, para tanto, o pacote *rbcb*, disponível no repositório oficial da rede CRAN (The Comprehensive R Archive Network) do software *R*, mediante o qual é feita a integração do *R* com as APIs (Application Programming Interface) – a ferramenta para extração de dados disponíveis na Internet – do SGS4 – Sistema Gerenciador de Séries Temporais do Banco Central do Brasil (Bacen).

2. Primeiros passos (instalar pacotes para a utilização da API e carregar na memória)

Antes de executar os exemplos é necessário instalar os pacotes das APIs para que as funções fiquem disponíveis para utilização, bem como, carregar o pacote na memória. Toda vez que uma nova sessão do software *R* for iniciada, o usuário deverá carregá-lo. Se isto não for feito, não será possível executar os exemplos constantes nesta nota técnica. Para carregar os pacotes necessários para os exemplos, utiliza-se, primeiramente, o pacote *pacman*, o qual servirá de base para carregar os demais, conforme segue:

```
if (!require("pacman")) install.packages("pacman")
pacman::p_load(rbcb, dplyr, ggplot2)
```

O pacote *pacman* possui ferramentas para executar tarefas associadas a pacotes complementares de maneira mais simplificada. A sua função `p_load()` verifica se os pacotes já estão instalados e, caso não tenham sido, os instala e carrega na memória. Solicitou-se o carregamento dos pacotes: (i) *rbcb*: utiliza as APIs do Bacen; (ii) *dplyr*: manipula os dados; e (iii) *ggplot2*: cria gráficos.

3. Pacote `rbcB` – SGS (Sistema Gerenciador de Séries Temporais do Banco Central do Brasil)

Os APIs do Bacen possibilitam o acesso a um conjunto de dados que dizem a respeito à economia brasileira, onde é possível encontrar dados como atividade econômica, economia regional, economia internacional e finanças públicas.

Criado pelo autor Wilson Freitas, o pacote `rbcB` está disponível na rede CRAN. Ele possui um conjunto amplo de funções para captação dos dados disponíveis no sistema SGS. Para exemplificar, utilizaremos duas de suas funções: a `get_series()` e a `get_corrency()`.

3.1. Função `get_series()`

Esta função, cuja sintaxe é apresentada a seguir, é utilizada para buscar variáveis de séries temporais disponíveis no SGS por meio do código da série, sendo possível também configurar o período desejado, conforme exemplifica o script apresentado a seguir:

Sintaxe

```
get_series(code, start_date = NULL, end_date = NULL, last = 0, as =  
c("tibble", "xts", "ts", "data.frame", "text"))
```

Exemplo 1 - Variação do PIB

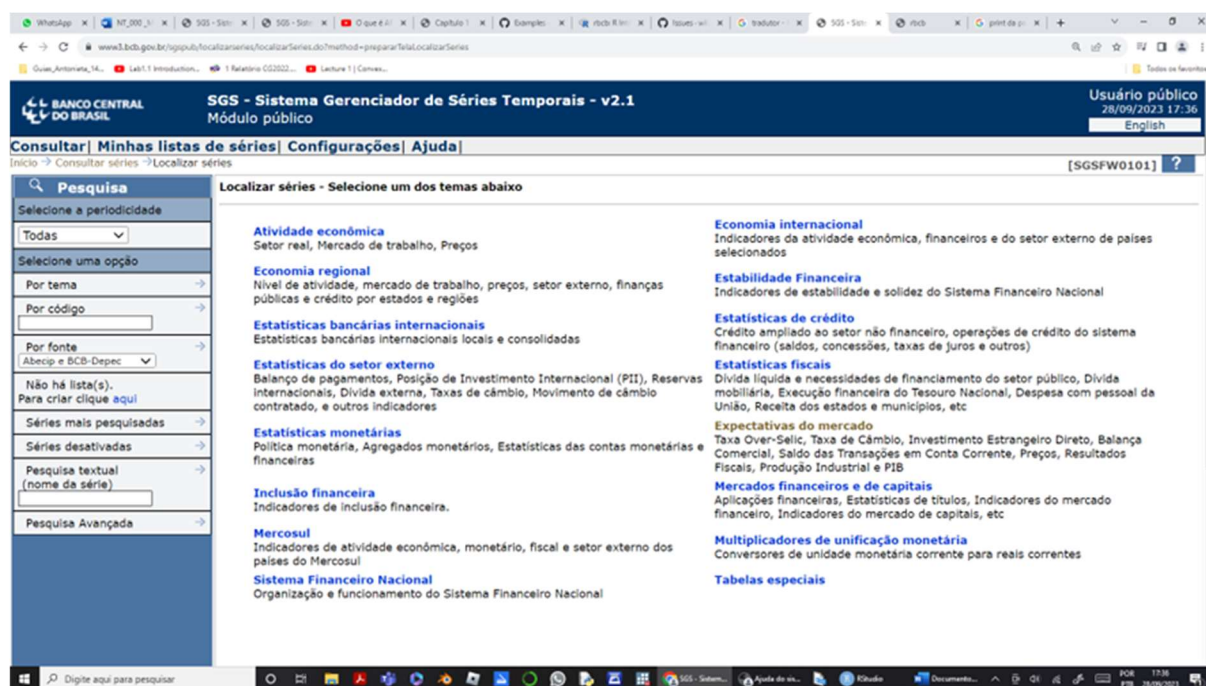
```
# Variação do PIB - Taxa de variação real no ano código = 7326  
  
PIB <- get_series(7326, start_date = '2010-01-01')  
head(PIB)  
  
>head(PIB)  
# A tibble: 6 × 2  
  date `7326`  
  <date> <dbl>  
1 2010-01-01 7.53  
2 2011-01-01 3.97  
3 2012-01-01 1.92  
4 2013-01-01 3  
5 2014-01-01 0.5  
6 2015-01-01 -3.55
```

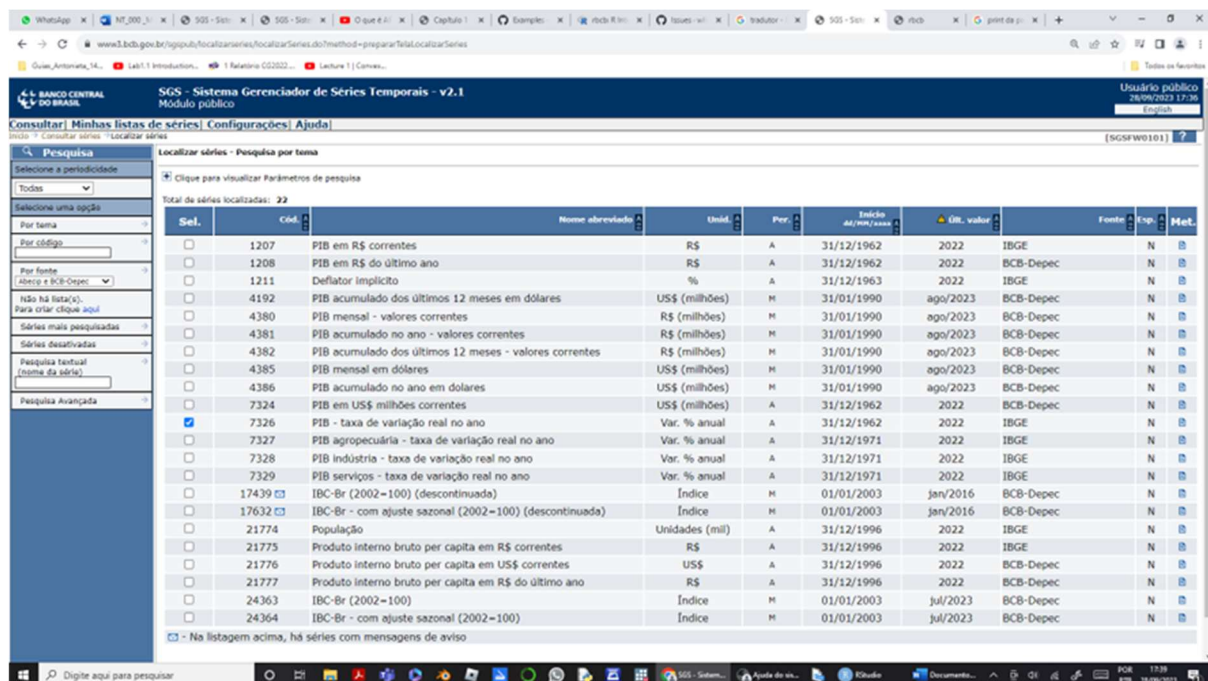
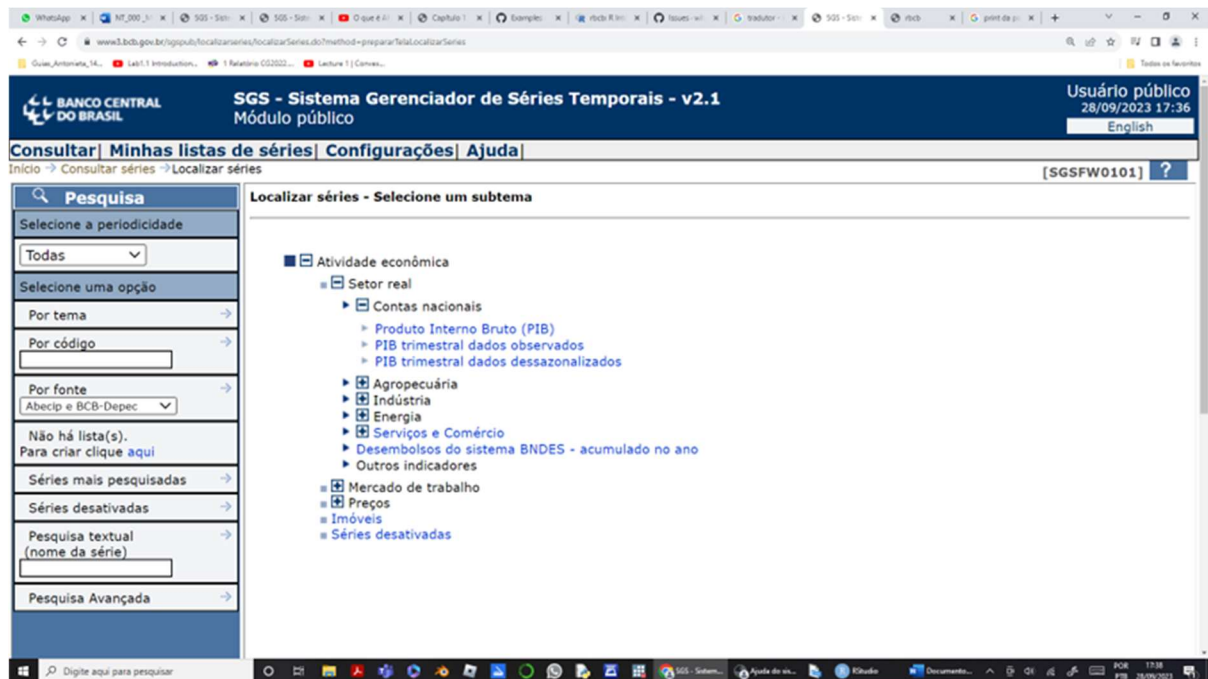
Neste exemplo, o código 7326 diz a respeito à série temporal “PIB – taxa de variação real no ano”. Os argumentos `start_date` e `end_date` delimitam, respectivamente, a data inicial e final da série, no formato americano “AAAA-MM-DD” (ANO-MÊS-DIA). Neste

caso, a série tem início em 1º de janeiro de 2010, mas não foi informada a data final, de modo que o último dado corresponde à última data registrada no sistema.

3.2 Consulta dos códigos

O código numérico da função `get_series()`, com o qual identifica-se a série temporal, é encontrado no SGS3, mediante seleção do tema de interesse, destacado em azul nos prints a seguir, conforme utilizado para a seleção do código 7326 do exemplo apresentado anteriormente.





3.3 Função `get_currency()`

Outra importante função do pacote é a `get_currency()`, cuja sintaxe é apresentada a seguir.

Sintaxe

```
get_currency(symbol, start_date, end_date, as = c("tibble", "xts", "data.frame", "text"))
```

Informando-se o símbolo da moeda e um intervalo de tempo (em datas), a função retorna a série temporal das taxas de câmbio de compra e venda, em Real (BRL). Para consultar a lista de moedas, utiliza-se o comando `list_currencies()`, que retorna nome, símbolo, código numérico, nome do país e código numérico do país.

Exemplo 2: lista de moedas

```
# Listar os códigos das moedas e dos países
> list_currencies()
# A tibble: 218 × 5
      name          code symbol country_name country_code
  <chr>      <dbl>  <chr>      <chr>      <dbl>
1 AFEGANE AFEGANIST      5      AFN      AFEGANISTAO      132
2 RANDE/AFRICA SUL     785      ZAR      AFRICA DO SUL     7560
3 LEK ALBANIA REP     490      ALL      ALBANIA, REPUBLICA DA 175
4 EURO             978      EUR      ALEMANHA          230
5 KWANZA/ANGOLA      635      AOA      ANGOLA           400
6 DOLAR CARIBE ORIENTAL 215      XCD      ANGUILLA          418
7 DOLAR CARIBE ORIENTAL 215      XCD      ANTIGUA E BARBUDA  434
8 RIAL/ARAB SAUDITA    820      SAR      ARABIA SAUDITA     531
9 DINAR ARGELINO       95      DZD      ARGELIA           590
10 PESO ARGENTINO      706      ARS      ARGENTINA         639
# i 208 more rows
# i Use `print(n = ...) ` to see more rows
```

Exemplo 3: taxas de câmbio

```
# get_currency - Taxa de câmbio dólar e euro
EUA <- get_currency('USD' , start_date = '2020-01-01', end_date = '2023-
09-23')

EUR <- get_currency('EUR' , start_date = '2020-01-01', end_date = '2023-
09-23')

moeda <- EUA %>%

# Renomear as colunas "ask" e "bid"
rename(compra_USD = 'ask', venda_USD = 'bid') %>%

# Juntar as bases pela coluna "date"
merge(EUR, by = 'date') %>%

# Renomear as colunas "ask" e "bid"
rename(compra_EUR = 'ask', venda_EUR = 'bid') %>%

# Selecionar apenas as colunas de preço de venda
select(date, venda_USD, venda_EUR)

# Ler as 5 primeiras linhas
head(moeda, 5)
      date      venda_USD      venda_EUR
1 2020-01-02      4.0207      4.5028
2 2020-01-03      4.0516      4.5224
3 2020-01-06      4.0548      4.5365
4 2020-01-07      4.0835      4.5494
5 2020-01-08      4.0666      4.5212
```

Conforme o exemplo, a função retorna 3 colunas, a data do período selecionado, a taxa de câmbio de compra (ask) e a taxa de câmbio de venda (bid). A função `get_currency()` não permite solicitar mais de uma moeda, então foram criadas duas tabelas, USD e EUR, manipuladas para se chegar a um *data frame* final, chamado *moeda*. Adicionalmente, filtrou-se apenas o valor médio de venda da taxa de câmbio das moedas.

4. Aplicação dos dados (Criação de gráfico em `ggplot`)

Após carregar as bases, pode-se criar gráficos por meio do pacote `ggplot`. Entretanto, o objetivo desta nota técnica não é ensinar o básico de `ggplot` e sim apresentar formas de aplicação dos dados obtidos via pacote `rbcbr`, bem como algumas soluções no `ggplot`.

Para exemplificar, produziremos 2 gráficos. No primeiro, usaremos a base da dados carregados no Exemplo 1, onde importamos a variação real do PIB anual no período de 2010 a 2022. Para este exemplo, criou-se um gráfico de barras usando a função `geom_bar()`, que representa uma melhor visualização para variação em anos, conforme script seguinte.

Exemplo 4: gráfico de barras da variação do PIB

```
# Gráfico de barras da variação do PIB
PIB %>%

# Criando uma coluna de rótulos para corrigir o problema de "," e "."
mutate(rotulo = gsub('[.]', ', ', as.character(`7326`))) %>% ggplot()

# Gráfico de barras definindo quem é a variável x e qual é y
geom_bar(aes(x = date, y = `7326`, fill = `7326` > 0),
  stat = "identity", position = position_dodge(preserve = "single"))

# Adicionando o rotulos de dados
geom_label(aes(x = date, y = `7326`, label = `7326`), stat = "identity",
  position = "identity")

# Escolhendo manualmente a cor das barras
scale_fill_manual(guide = FALSE, breaks = c(TRUE, FALSE),
  values=c("#29B6A5", "#b6293a"))

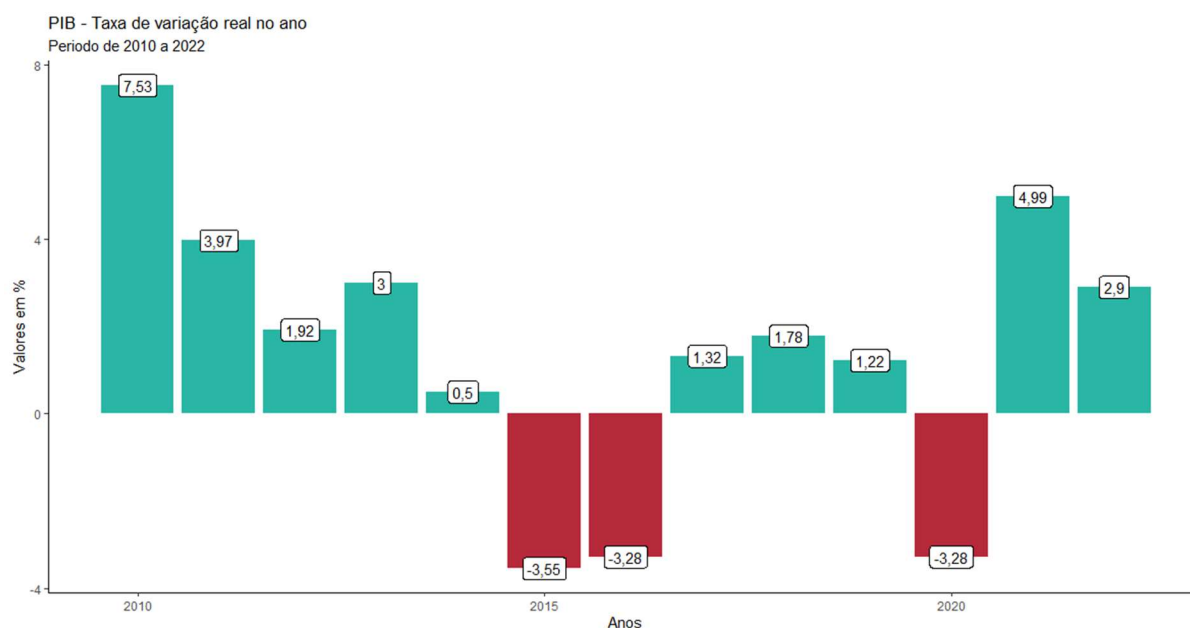
# Titulos e subtítulos
labs(x = "Anos", y = "Valores em %", title = "PIB - Taxa de variação real
no ano", subtitle = "Período de 2010 a 2022", linetype = "Variable",
  color = "Variable")

# Alterando a estetica
theme_classic()
```

Para corrigir um problema de rótulos de dados por conta da diferença entre formatação de números americana e formatação brasileira, criou-se uma coluna fantasma com a função `mutate()` do pacote `dplyr`, bem como uma coluna que substituirá os pontos, “.”, por vírgulas, “,”. Adicionalmente, e tendo em vista que em certos anos houve variação negativa do PIB no Brasil, foi incorporada uma condição no `geom_bar()`, especificamente no comando `fill`. Dessa forma, se o valor do dado for maior que zero ele retorna `TRUE` ou `FALSE`. Em seguida, para a escolha manual da cor das barras, usa-se a função `scale_fill_manual()`, onde podemos identificar pelo comando `breaks` quais são as colunas que devem ser coloridas e pelo comando `values` a lista de cor desejada.

Por fim, inclui-se a função `geom_label()` para inserir os rótulos de dados em cada barra. É possível utilizar a função `geom_text()`, que possui uma função parecida, porém é mais trabalhoso para incorporar no gráfico. Segue o resultado dos dados apresentados:

Gráfico 1: PIB – Taxa de variação real por ano



Fonte: Bacen. Elaborado pelos autores

Para a elaboração do gráfico conforme o conjunto de dados do Exemplo 3, relativo ao preço médio de câmbio, utiliza-se o gráfico de linhas, o qual é o mais indicado para quando há dados em série temporal. O principal comando no `ggplot` para a execução é o `geom_line()`. É importante mencionar que no pacote `ggplot` é possível criar gráficos em camadas, conforme script apresentado a seguir, mediante o qual foi elaborado o Gráfico 2.

Exemplo 5: Criação do gráfico em linhas e por camadas do preço médio de venda do Dólar e do Euro

```
# Chamando os dados e o ggplot
moeda %>% ggplot()

# Primeira camada - linhas da moeda USD
geom_line(aes(x = date, y = venda_USD, color="USD"), size = 1)

# Segunda camada - linhas da moeda EUR
geom_line(aes(x = date, y = venda_EUR, color="EUR"), size = 1)

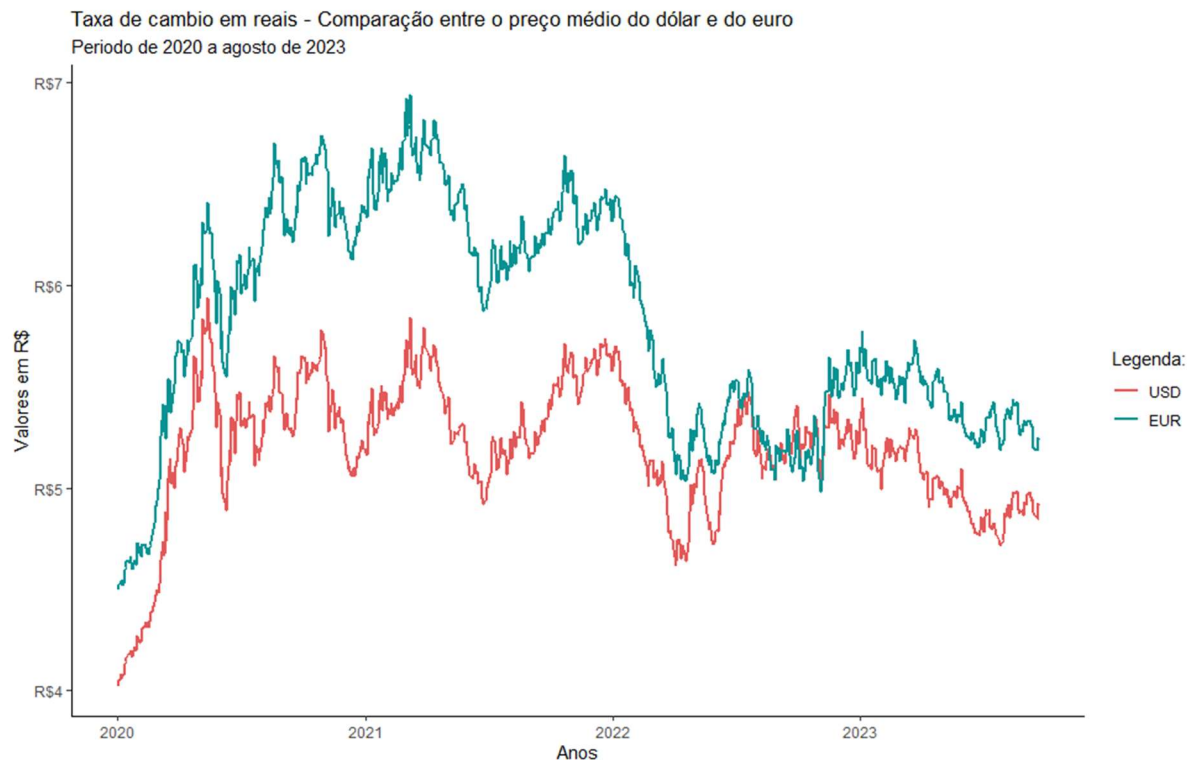
# Inserir o prefixo da moeda R$ no eixo y
scale_y_continuous(labels=scales::label_number(prefix = "R$"))

labs(x = "Anos", y = "Valores em R$", title = "Taxa de câmbio
valor de venda", subtitle = "Período de 2010 a agosto de
2023", linetype = "Variable", color = "Variable")

# Ajustar manualmente a legenda por conta de mais de uma
camada
scale_color_manual(breaks = c('USD', "EUR"), values = c('USD'
= '#e35959', "EUR" = "#099393"), name="Legenda:")

theme_classic()
```

Gráfico 2: Taxa de câmbio em reais – preço médio do dólar versus euro



Fonte: Bacen. Elaborado pelos autores

Apesar de haver muito exemplos na Internet de gráficos em linhas elaborados com o `ggplot`, há poucos que mencionam a possibilidade de fazer um gráfico multicamadas, ou seja, usar coluna a coluna para traçar a linha dos dados.

5. Considerações finais

Este trabalho é uma introdução ao pacote `rbcb` da linguagem *R*, utilizando-se como exemplos de aplicações as funções `get_series()` e `get_currency()`, bem como algumas funções dos pacotes `ggplot` e `dplyr`. Registre-se que os citados pacotes possuem várias outras funções, de modo que apenas uma pequena parte do assunto foi abordada, suficiente, contudo, para a confecção de estudos iniciais e, enfim, a possibilidade do usuário buscar o aprofundamento nas inúmeras publicações sobre o tema, especialmente no próprio CRAN.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Banco Central do Brasil. *Sistema Gerenciador de Séries Temporais do Banco Central do Brasil*. Disponível em: <<https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>>. Acesso em: 28 set. 2023.
- [2] R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>
- [3] *R Interface to Brazilian Central Bank Web Services*. Disponível em: <<https://cran.r-project.org/web/packages/rbcbrbcbr.pdf>>. Acesso em: 28 set. 2023.