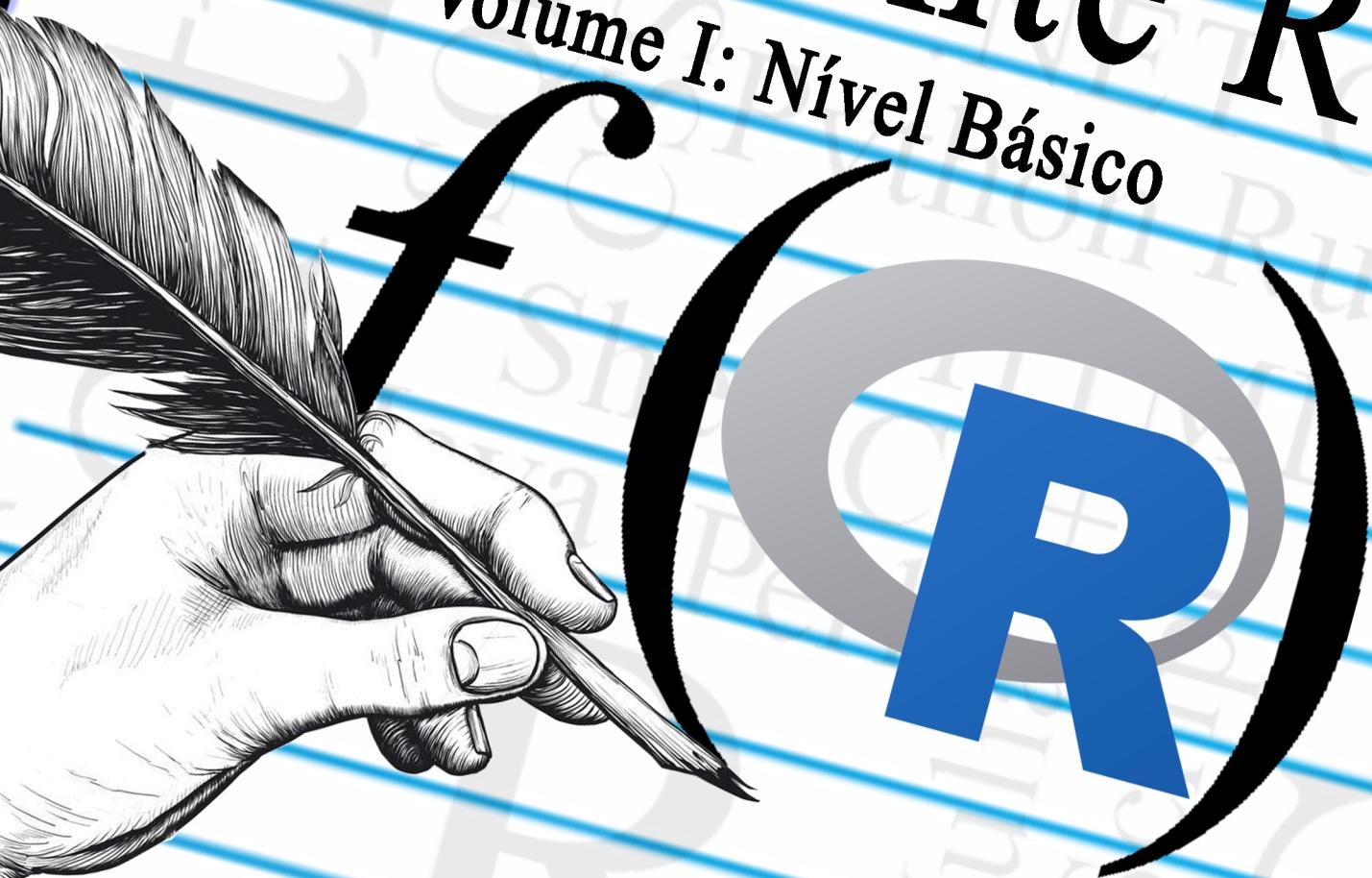


# Estudando o Ambiente R

Volume I: Nível Básico



Ben Dêivide  
Diego Arthur



Democratizando  
Conhecimento



# **ESTUDANDO O AMBIENTE R**

**BEN DÊIVIDE DE OLIVEIRA BATISTA  
DIEGO ARTHUR BISPO JUSTINO DE OLIVEIRA**



# **Estudando o Ambiente R**

*Volume I: Nível Basico*

BEN DÊIVIDE DE OLIVEIRA BATISTA  
DIEGO ARTHUR BISPO JUSTINO DE OLIVEIRA



Democratizando  
Conhecimento

Ouro Branco, MG, 25 de outubro de 2021

© 2021 by Ben Dêivide de Oliveira Batista e Diego Arthur Bispo Justino de Oliveira



Esse material está licenciado com uma Licença Creative Commons - Atribuição - Não Comercial 4.0 Internacional. Usamos também a filosofia de trabalho com o Selo Democratizando Conhecimento (DC). O leitor é livre para compartilhar, redistribuir, transformar ou adaptar essa obra, desde que não venha a utilizá-la em nenhuma atividade de propósito comercial. Por fim, a única exigência é a atribuição dos dos créditos aos autores dessa obra.

Direitos de publicação reservados ao seu conhecimento.

Impresso no Brasil - ISBN (Digital):

Impresso no Brasil - ISBN (Impresso):

**Projeto Gráfico:** Ben Dêivide de Oliveira Batista

**Revisão técnica e textual:** XX

**Revisão de Referências Bibliográficas:**

**Editoração Eletrônica:** Ben Dêivide de Oliveira Batista

**Capa:** Ben Dêivide de Oliveira Batista

**Como citar essa obra:**

BATISTA, B. D. O.. **Estudando o Ambiente R.** 1ed. Ouro Branco, MG:[sn]. 2021. 1 v. Disponível em: <<https://bendeivide.github.io/book-eambr02/>>

**Autor correspondente e mantenedor da obra:**

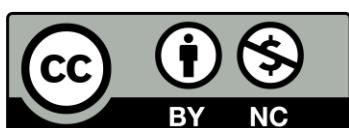
Ben Dêivide de Oliveira Batista

Contato: <[ben.deivide@gmail.com](mailto:ben.deivide@gmail.com)>

Site pessoal: <<http://bendeivide.github.io/>>

# Licença

Todos os direitos autorais contidos nesse livro são reservados ao seu conhecimento, usufrua-o, pois é totalmente de graça. Use-o com responsabilidade e saiba valorizar.



Este trabalho está licenciado com uma Licença Creative Commons - Atribuição - Não Comercial 4.0 Internacional. Usamos também a filosofia de trabalho com o Selo Democratizando Conhecimento (DC).





## Epígrafe

*A melhor linguagem é a que você domina!*  
**Ben Dêivide**

# Sumário

<b>Licença</b>	<b>i</b>
<b>Epígrafe</b>	<b>iii</b>
<b>Prefácio</b>	<b>v</b>
<b>1 Entendendo a coleção Estudando o ambiente R</b>	<b>1</b>
1.1 Volume I: Nível Básico . . . . .	2
1.2 Volume II: Nível Intermediário . . . . .	2
1.3 Volume III: Nível Avançado . . . . .	2
1.4 Demais volumes . . . . .	3
1.5 Referências complementares da Coleção . . . . .	3
1.6 Exemplos . . . . .	4
Exercícios propostos . . . . .	5
<b>2 História e instalação do R</b>	<b>6</b>
2.1 História do R . . . . .	6
<b>3 Medidas de Posição</b>	<b>7</b>
3.1 Introdução . . . . .	7
<b>4 Gabarito dos Exercícios</b>	<b>8</b>

# Prefácio

A coleção *Estudando o ambiente R* é fruto de cursos ministrados sobre essa linguagem, bem como consultorias e estudos ao longo dos anos. Em 2005, quando ingressei na academia no curso de Engenharia Agronômica fiquei fascinado com a disciplina de Estatística no segundo semestre do ano corrente. Na sequência, acabo tendo o primeiro contato com o ambiente R, com pouco mais de 9 anos de seu lançamento e redistribuição. Poucos materiais naquela época haviam disponíveis em língua portuguesa. Porém, foi o suficiente para eu entender que estava diante de uma grande ferramenta computacional e estatística, necessária para o entendimento, pois sabia que poderia me gerar além de conhecimento, bons frutos acadêmicos.

Hoje, no ano de 2021, usuário há mais de 15 anos dessa linguagem, percebi que me sentia desconfortável, como apenas usuário dessa ferramenta de trabalho. E assim, quando queremos aprender algo não há ferramenta melhor do que *\*aprender por ensinar\**. E assim, lotado no Departamento de Estatística, Física e Matemática (DEFIM), campus Alto Paraopeba, pela Universidade Federal de São João del-Rei (UFSJ), juntamente com o Centro Acadêmico de Engenharia de Telecomunicações (UFSJ), resolvemos em parceria, ministrar nesse momento de pandemia uma sequência de módulos para o curso R, desde o nível Básico até ao módulo Avançado.

A ideia desse curso foi apresentar algo diferente relacionado a maioria dos cursos em R, que foi sempre apresentar essa ferramenta dentro dos conceitos da área da Estatística. Apesar de uma coisa ser intrínseca a outra, há muitas particularidades no ambiente R que são complexos, e muitas vezes julgados erroneamente. Um dos exemplos clássicos é que *loops* em R são lentos e com alto gasto de memória, quando na realidade, isso ocorre muitas vezes pelo não entendimento do sistema de cópia de objetos nesse ambiente. Ainda mais, o entendimento desses cursos é agravado porque o entendimento sobre a estatística além de um cunho matemático, tem o seu cunho filosófico de como as metodologias foram desenvolvidas, e o entendimento mútuo da Estatística e o ambiente R, podem não ter o conhecimento real que essa potencial ferramenta pode proporcionar, uma vez que muitos assuntos complexos podem estar envolvidos em uma única aula.

Assim, desenvolvemos na coleção *Estudando o ambiente R* os três volumes iniciais, referentes a apenas a linguagem R, sendo *Volume I: Nível Básico*, *Volume II: Nível Intermediário* e *Volume III: Nível Avançado*. Fazendo a alusão dos três livros iniciais sobre a linguagem S de John Chambers, faremos uma explanação sobre assuntos de menor complexidade até noções mais complexas sobre o ambiente R, restringindo apenas a sintaxe e semântica da linguagem. Os volumes subsequentes serão destinados a *Documentações no R*, *Desenvolvimento de pacote R*, *Gráficos*, *Banco de dados*, *Interface Gráfica ao Usuário*, *Interface R com outras linguagens*, dentre outros.

Tentando engajar nossos alunos, e agora colegas de trabalho, tenho a parceria no Volume I, de Diego Arthur, uma pessoa que tenta se superar a cada desafio e assunto estudado.

Por fim, espero que esse primeiro volume possa servir de referência para os passos iniciais nessa ferramenta tão importante para a área de análise de dados.

**Ben Dêivide de Oliveira Batista**  
Ouro Branco, MG, 25 de outubro de 2021



## *Entendendo a coleção Estudando o ambiente R*

A Coleção *Estudando o ambiente R* não tem como objetivo principal de ensinar análise de dados. Mas sim, proporcionar ao leitor um conhecimento sobre a linguagem **R**, de modo que se possa usufruir todos os recursos que esse ambiente possa proporcionar.

Ainda como complemento, não queremos nesse material, convencê-lo a utilizar a linguagem **R**, pois a melhor linguagem é aquela que você domina. Contudo, pretendemos mostrar que os recursos utilizados pelo **R** não estão mais limitados a própria análise de dados. Um exemplo é o material didático dessa coleção que pode ser acessada em: <<https://bendeivide.github.io/cursor/>>, que nesse momento usufruímos da própria linguagem para repassar as nossas experiências sem ao menos ter o domínio sobre linguagens do tipo HTML, CSS, JavaScript, dentre outras, necessárias para uma boa renderização de uma página *web*. Isso mostra a potencial ferramenta de trabalho que o ambiente **R** pode ser para a vida profissional.

Dessa forma, propormos um entendimento sobre a sintaxe e semântica de como a linguagem **R** é desenvolvida. Com isso, o leitor será capaz após a leitura dos dois primeiros volumes, de acompanhar qualquer curso de Estatística com aplicações em **R**, se dedicando apenas ao entendimento na área da Estatística, uma vez que o embasamento sobre o ambiente **R** foi suprido por essa coleção. Essa nova revolução do Dados, se deve ao grande volume de informações obtidos nessa era tecnológica. Juntamente com a Estatística, o **R** se tornará uma poderosa ferramenta para entender os padrões que estão por trás dos dados, que por sinal, é a moeda valiosa do momento, ou melhor, sempre foi!

Aprenderemos também recursos diversos na área da computação, como programação defensiva, desenvolvimento de interfaces gráficas, paralelização, como também recursos na área da estatística sem complexidades teóricas, como o desenvolvimento de gráficos e o uso de banco de dados. Ensinaremos também o desenvolvimento de materiais como artigos, livros, *websites*, *blogs*, *dashboards*. Por fim, chegaremos a maior cobiça de um programador **R**, desenvolver um pacote.

### Por que os artigos “o” e “a” para o **R**?



Observem que em muitos momentos utilizamos o artigo “o” para a linguagem **R**. Pois é, isso ocorre porque ela também é considerada um software ou ambiente. Daí, também podemos chamá-la de software **R**, ou preferivelmente, ambiente **R**.

Os módulos dessa coleção terão os três volumes base para o entendimento do ambiente **R**:

- Volume I: Nível Básico;
- Volume II: Nível Intermediário; e
- Volume III: Nível Avançado.

A seguir, explanaremos sobre cada um dos módulos.

## 1.1 Volume I: Nível Básico

Esse primeiro volume, que representa o livro corrente, apresenta um breve **histórico** sobre a linguagem, a sua instalação, bem como os recursos que a **IDE<sup>1</sup> RStudio**, o conhecimento da **sintaxe** e **semântica** da linguagem **R**, compreendendo as estruturas bases da linguagem, sobre o que é um **objeto** e como construir uma **função**, o entendimento sobre **fluxos de controle**. O que é um **pacote**, **carregar** e **anexar** um pacote, e quem são as pessoas que fazem parte da manutenção dessa linguagem, também serão assuntos desse primeiro módulo. **Caminho de busca, ambientes** e *namespaces*, teremos noções básicas. Algo muito interessante, que pode mudar a vida de um programador em **R** são as **boas práticas para a escrita de um código**, tema também abordado nesse módulo.

A ideia desse volume é proporcionar um entendimento básico, um primeiro contato com a linguagem, fazendo com que o leitor possa dar os primeiros passos, executando as primeiras linhas de comando. Mas também, dando o enfoque com erros tão recorrentes, como o entendimento sobre um objeto, ou o anexo de um pacote no caminho de busca. Temas como esses, dentre outros, serão a forma inicial que encontramos, para que posteriormente, seja dado um aprofundamento sobre a estrutura de um objeto **R** bem como a sua manipulação, e adicionado a isso, a inserção de como são os paradigmas da programação nesse ambiente. Essa última parte será estudada, no Volume II, apresentado a seguir.

## 1.2 Volume II: Nível Intermediário

O volume II é introduzido com uma melhor caracterização do ambiente **R** quanto ao seu **escopo léxico**, como **linguagem interpretada**, como **programação funcional**, como **programação meta-paradigma**, como **programação dinâmica**; apresentaremos **manipulações de objetos em mais detalhe**, bem como o surgimento de alguns outros objetos como **tibble**, **cópias de objetos**. Uma característica do ambiente **R** é que a linguagem pode ser **orientada a objetos** e isso será estudado nesse módulo. Introduziremos ao **desenvolvimento de pacotes R**, e aprofundaremos sobre os **ambientes**. Por fim, mostraremos como desenvolver Projeto do **RStudio** e integrá-los ao **GitHub**, e dessa forma, introduziremos sobre o sistema **Git**.

Esse talvez seja o maior volume, dentre os três iniciais, porque apesar de não precisarmos entender mais a ideia dos objetos, que foram retratadas no Volume I, a inserção dos paradigmas da programação para este volume, trará uma maior riqueza de características para o **R**, mostrando a sua versatilidade. Também, daremos um maior detalhamento como manipular objetos, e as otimizações existentes da linguagem, como por exemplo, a modificação no local, que se entendida, poderá perceber que o **loop** no ambiente **R** não é lento quanto parece. Ao final desse volume, falaremos sobre como propagar o seu código com o sistema **Git** na plataforma **GitHub**, sincronizado com os projetos do **RStudio**.

## 1.3 Volume III: Nível Avançado

O Volume III, será a total exploração do manual **R Internals**. Apesar de ser um assunto voltado para membros do **R Core Team**, pretendemos entender como o **R** trabalha nos bastidores. Dessa forma, teremos total controle sobre as nossas rotinas. Contudo, para usuários que pretendem entender o ambiente **R** de forma aplicada, pode avançar esse volume para a leitura dos volumes seguintes.

Nesse volume, faremos uma introdução sobre a linguagem **C**, e entender algumas estruturas, como por exemplo ponteiros, e instruções, como por exemplo, **switch()**, e perceber que a arquitetura dos objetos em **R**, são desenvolvidas dentro dessas ideias. Também será possível usar a linguagem **C** sem necessidade de pacote adicional. Faremos também uma introdução sobre a linguagem **FORTRAN** e **S**, as duas outras linguagens complementares para o entendimento completo dos bastidores do **R**.

<sup>1</sup>Do inglês, *Integrated Development Environment*, que significa ambiente de desenvolvimento integrado.

## 1.4 Demais volumes

Os demais volumes compreendem lacunas necessárias para serem abordadas com profundidade, tais como: **Documentações no R**, **Desenvolvimento de pacote R**, **Gráficos**, **Banco de dados**, **Interface Gráfica ao Usuário**, **Interface R com outras linguagens**, **Funções do pacote base**, dentre outros.

## 1.5 Referências complementares da Coleção

Citaremos alguns livros e materiais utilizados para o desenvolvimento dessa coleção, que alguns podem ser acessados *online*, como também via **bookdown**, tais como:

- Manuais do **R** :
  - An Introduction to R (??);
  - R Data Import/Export (??);
  - R Installation and Administration (??);
  - Writing R extensions (??);
  - R language definition (??);
  - R Internals (??);
- *Bookdowns*:
  - Advanced R (??);
  - Advanced R Solutions (??);
  - R Packages (??);
  - R for Data Science (??);
  - Hands-On Programming with R (??);
  - R Markdown (??);
  - bookdown (??);
  - Dynamic documents with R and knitr (??);
- Livros físicos:
  -

## 1.6 Exemplos

### Código R 1.1

Script:

```
1 # BP Statistical Review 2018
2
3 # Producao e consumo de petroleo no brasil (Milhares de barris por dia)
4 ano <- as.factor(c(2007:2017, 2007:2017))
5 prodcons <- c(1831, 1897, 2029, 2137, 2179, 2145, 2110, 2341, 2525, 2608,
   2734, 2308, 2481, 2498, 2716, 2839, 2915, 3124, 3242, 3181, 3013, 3017)
6 id <- c(rep("Produção", 11), rep("Consumo", 11))
7
8 # Objeto que armazena as informaçoes
9 dados <- data.frame(ano, Legenda = id, prodcons)
10
11 # Pacote utilizado
12 install.packages("ggplot2")
13 library(ggplot2)
14
15 # Funcao para criacao do grafico de barras
16 ggplot(dados, aes(x=ano, y=prodcons, fill=Legenda)) +
17   geom_bar(position="dodge", stat="identity") +
18   xlab("Ano") + ylab(``Petróleo (Milhões de barris por dia)'')
```

## Exercícios propostos

Apresentamos alguns exercícios, dos quais se conseguir resolvê-los entendendo a sua complexidade, poderá avançar para a leitura do Volume II.

**Exercício 1.1:** Texto ...

*Solução na página 8*

## *História e instalação do R*

### 2.1 História do R



## *Medidas de Posição*

### **3.1 Introdução**



# Gabarito dos Exercícios

## Solução dos Exercícios do Capítulo 1

Solução do Exercício 1.1 na página 5:  
Solução ...



---

## **Solução dos Exercícios do Capítulo 2**

### Solução dos Exercícios do Capítulo 3

# Referências Bibliográficas

- GROLEMUND, G. *Hands-On Programming with R: Write Your Own Functions and Simulations*. Sebastopol: O'Reilly Media, 2014. ISBN 9781449359119. Disponível em: <<https://rstudio-education.github.io/hopr/>>.
- GROSSER, M.; BUMAN, H.; WICKHAM, H. *Advanced R Solutions*. 2nd. ed. Boca Raton, Florida: Chapman and Hall/CRC, 2021. ISBN 9781000409079. Disponível em: <<https://adv-r.hadley.nz/>>.
- R CORE TEAM. *R data Import/Export*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- R CORE TEAM. *R instalation and administration*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- R CORE TEAM. *R internals*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- R CORE TEAM. *R language definition*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- R CORE TEAM. *Writing R extensions*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- VENABLES, W. N.; SMITH, D. M.; R CORE TEAM. *An Introduction to R: Notes on R a programming environment for data analysis and graphics*. Vienna, Austria, 2021. Version 4.1.1 (10-08-2021). Disponível em: <<https://cran.r-project.org/manuals.html>>.
- WICKHAM, H. *R Packages*. 2nd. ed. Sebastopol: O'Reilly Media, 2015. ISBN 9781491910597. Disponível em: <<https://r-pkgs.org/index.html>>.
- WICKHAM, H. *Advanced R*. 2nd. ed. Boca Raton, Florida: Chapman and Hall/CRC, 2019. ISBN 978-0815384571. Disponível em: <<https://adv-r.hadley.nz/>>.
- WICKHAM, H.; GROLEMUND, G. *R for Data Science*. Sebastopol: O'Reilly Media, 2017. ISBN 978-1-491-91039-9. Disponível em: <<https://r4ds.had.co.nz/>>.
- XIE, Y. *Dynamic documents with R and knitr*. 2. ed. Boca Raton, Florida: Chapman and Hall/CRC, 2015. ISBN 978-1498716963. Disponível em: <<https://yihui.org/knitr/>>.
- XIE, Y. *bookdown: Authoring books and technical documents with r markdown*. Boca Raton, Florida: Chapman and Hall/CRC, 2017. ISBN 978-1138700109. Disponível em: <<https://bookdown.org/yihui/bookdown/>>.



