

# Introdução ao Tratamento e Análise de Dados em R

## Aula 1 - Instalando e Configurando o R e RStudio

Sérgio Rivero

PPGE-UFPA

12 de maio de 2019



# Sumário

- 1 Objetivos do Curso
- 2 Conteúdo da Aula
- 3 Exercícios

- 4 Dúvidas
- 5 Agradecimentos
- 6 Referências



# Objetivos do Curso

Neste curso queremos:

- 1 Apresentar os conceitos e estratégias básicas para o tratamento de dados em R
- 2 Apresentar as principais ferramentas de tratamento de dados do R (e seu ecossistema)
- 3 Aprender a integrar estas estratégias e ferramentas para produzir documentos e relatórios

Para isso utilizaremos uma estratégia onde:

- 50% do tempo será de apresentações sobre o R pelo instrutor
- 50% do tempo será de prática



# O que é o R?

- O R é uma suíte integrada de software que permite a recuperação, o tratamento, e a análise de dados[VS11].
- Pode se dizer que o R é um ambiente de tratamento de dados que permite ao usuário, além a análise de dados propriamente dita, escrever extensões e ampliar o seu escopo.
- R é uma ferramenta de software livre que atende aos critérios da *Free Software Foundation* e tem uma licença **GNU** - <https://www.gnu.org/>



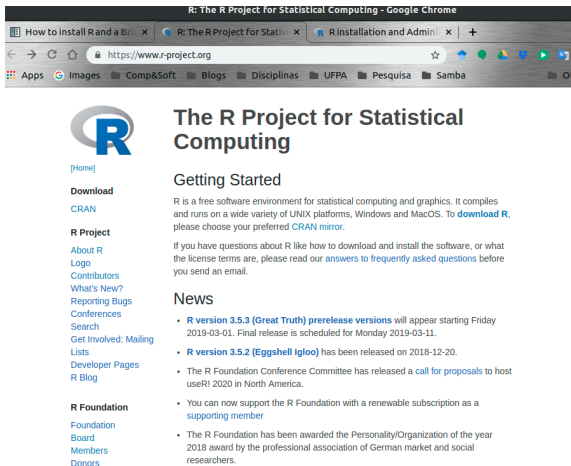
# Funcionalidades do R

Algumas das funcionalidade do R são[VS11]:

- Ferramentas para manuseio e armazenamento de dados
- Um conjunto de operadores que permitem o cálculo numérico e a manipulação de matrizes
- Um enorme conjunto de bibliotecas para análise de dados
- Ferramentas para apresentação gráfica de dados e resultados
- Uma linguagem de programação orientada a objetos e extensível
- A possibilidade de estender a linguagem, suas bibliotecas e funções



# Instalando o R: selecione **download R**



The screenshot shows the R Project for Statistical Computing website in a Google Chrome browser. The browser tabs include 'How to install R and a B...', 'R: The R Project for Statis...', and 'R Installation and Admini...'. The address bar shows 'https://www.r-project.org'. The website features the R logo, a navigation menu on the left, and main content sections for 'Getting Started' and 'News'.

**The R Project for Statistical Computing**

[Home]

**Download**

[CRAN](#)

**R Project**

[About R](#)  
[Logo](#)  
[Contributors](#)  
[What's New?](#)  
[Reporting Bugs](#)  
[Conferences](#)  
[Search](#)  
[Get Involved: Mailing Lists](#)  
[Developer Pages](#)  
[R Blog](#)

**R Foundation**

[Foundation](#)  
[Board](#)  
[Members](#)  
[Donors](#)

## Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To [download R](#), please choose your preferred [CRAN mirror](#).

If you have questions about R like how to download and install the software, or what the license terms are, please read our [answers to frequently asked questions](#) before you send an email.

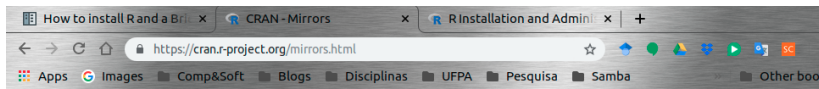
## News

- [R version 3.5.3 \(Great Truth\) prerelease versions](#) will appear starting Friday 2019-03-01. Final release is scheduled for Monday 2019-03-11.
- [R version 3.5.2 \(Eggshell Igloo\)](#) has been released on 2018-12-20.
- The R Foundation Conference Committee has released a [call for proposals](#) to host useR! 2020 in North America.
- You can now support the R Foundation with a renewable subscription as a [supporting member](#)
- The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.

Nesta página clicar em **download R**



# Instalando o R: selecione o espelho



## CRAN Mirrors

The Comprehensive R Archive Network is available at the following URLs, please choose a location close to you. Some statistics on the status of the mirrors can be found here: [main page](#), [windows release](#), [windows old release](#).

If you want to host a new mirror at your institution, please have a look at the [CRAN Mirror HOWTO](#).

### 0-Cloud

<https://cloud.r-project.org/>

Automatic redirection to servers worldwide, currently sponsored by Rstudio

<http://cloud.r-project.org/>

Automatic redirection to servers worldwide, currently sponsored by Rstudio

### Algeria

<https://cran.usthb.dz/>

University of Science and Technology Houari Boumediene

<http://cran.usthb.dz/>

University of Science and Technology Houari Boumediene

### Argentina

<http://mirror.fcaglp.unlp.edu.ar/CRAN/>

Universidad Nacional de La Plata

### Australia

<https://cran.csiro.au/>

CSIRO

<http://cran.csiro.au/>

CSIRO

<http://mirror.aarnet.edu.au/pub/CRAN/>

AARNET

<https://cran.ms.unimelb.edu.au/>

School of Mathematics and Statistics, University of Melbourne

<https://cran.curtin.edu.au/>

Curtin University of Technology

### Austria

<https://cran.wu.ac.at/>

Wirtschaftsuniversität Wien

<http://cran.wu.ac.at/>

Wirtschaftsuniversität Wien

### Belgium

<https://www.freeststatistics.org/cran/>

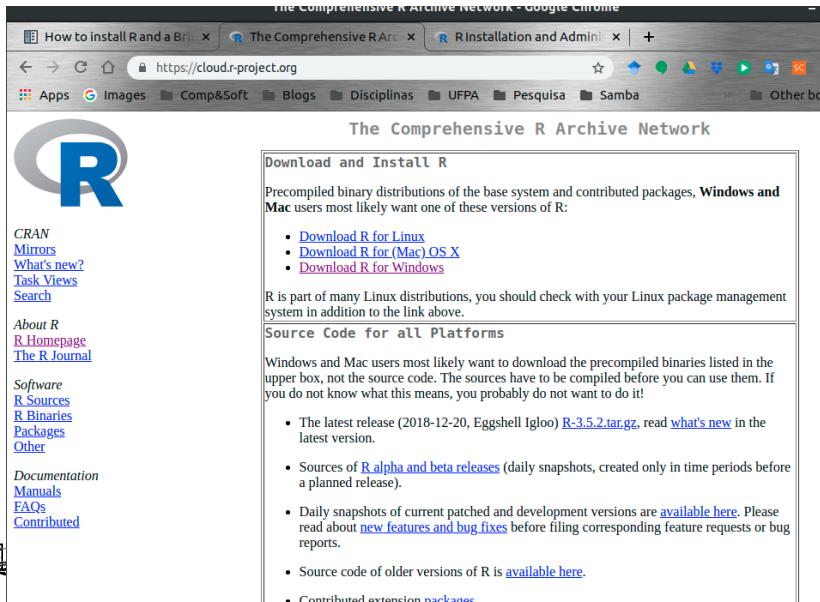
Patrick Wessa

<http://www.freeststatistics.org/cran/>

Patrick Wessa



# Instalando o R: baixe de Acordo com seu S.O.



The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (2018-12-20, Eggshell Igloo) [R-3.5.2.tar.gz](#), read [what's new](#) in the latest version.
- Sources of [R alpha and beta releases](#) (daily snapshots, created only in time periods before a planned release).
- Daily snapshots of current patched and development versions are [available here](#). Please read about [new features and bug fixes](#) before filing corresponding feature requests or bug reports.
- Source code of older versions of R is [available here](#).
- Contributed extension packages



# Instalando o R: aqui para o Windows

## R-3.5.2 for Windows (32/64 bit)

[Download R 3.5.2 for Windows](#) (79 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

### Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Página do Download do R



# Mais Informações sobre a Instalação do R

Cheque os *FAQs - Frequently Asked Questions*

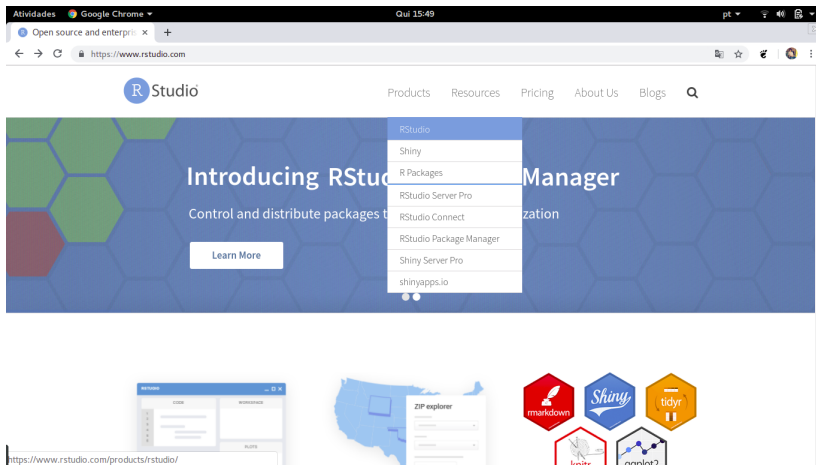
`https:`

`//cloud.r-project.org/bin/windows/base/rw-FAQ.html`

`https://cran.r-project.org/doc/manuals/R-admin.html`



# Instalando o RStudio



<https://www.rstudio.com/products/rstudio/>

<https://www.rstudio.com/>



# Aba de Produtos do RStudio

← → ↻ 🔒 <https://www.rstudio.com/products/rstudio/>

[Products](#)[Resources](#)[Pricing](#)[About Us](#)[Blogs](#)

## RStudio

### Take control of your R code

RStudio is an integrated development environment (IDE) for R. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, debugging and workspace management. [Click here to see more RStudio features.](#)

RStudio is available in open source and commercial editions and runs on the desktop (Windows, Mac, and Linux) or in a browser connected to RStudio Server or RStudio Server Pro (Debian/Ubuntu, RedHat/CentOS, and SUSE Linux).



#### Desktop

Run RStudio on your desktop

[RStudio Desktop >](#)



#### Server

Centralize access and computation

[RStudio Server >](#)



# Selecione seu RStudio

[Products](#)[Resources](#)[Pricing](#)[About Us](#)[Blogs](#)

## RStudio Desktop

### Open Source Edition

### Commercial License

#### Overview

- Access RStudio locally
- Syntax highlighting, code completion, and smart indentation
- Execute R code directly from the source editor
- Quickly jump to function definitions
- Easily manage multiple working directories using projects
- Integrated R help and documentation
- Interactive debugger to diagnose and fix errors quickly
- Extensive package development tools

All of the features of open source; plus:

- A commercial license for organizations not able to use AGPL software
- Access to priority support

#### Support

Community forums only

- Priority Email Support
- 8 hour response during business hours (ET)

#### License

AGPL v3

[RStudio License Agreement](#)

#### Pricing

Free

\$995/year

[DOWNLOAD RSTUDIO DESKTOP](#)[BUY NOW](#)

# Selecionando o "Plano"

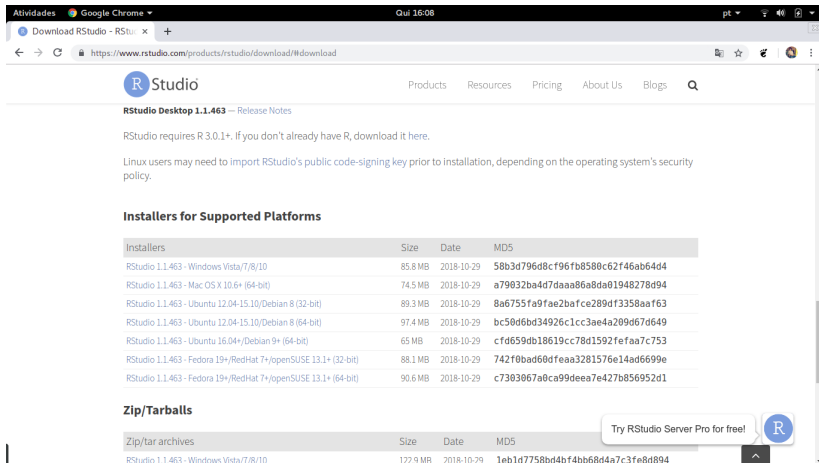
	RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License	RStudio Server Pro + RStudio Connect Commercial License
	FREE	\$995 per year	FREE	\$9,995 per year	\$29,995 per year
	<a href="#">DOWNLOAD</a> Learn More	<a href="#">BUY</a> Learn More	<a href="#">DOWNLOAD</a> Learn More	<a href="#">DOWNLOAD</a> Learn More	<a href="#">TALK</a> Learn More
Integrated Tools for R	●	●	●	●	●
Priority Support		●		●	●
Access via Web Browser			●	●	●
Enterprise Security				●	●
Project Sharing				●	●
Manage Multiple R Sessions & Versions				●	

Try RStudio Server Pro for free!

Planos, preços e benefícios



# Finalmente, instalando seu executável



**RStudio Desktop 1.1.463 — Release Notes**

RStudio requires R 3.0.1+. If you don't already have R, download it [here](#).

Linux users may need to import RStudio's [public code-signing key](#) prior to installation, depending on the operating system's security policy.

### Installers for Supported Platforms

Installers	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	85.8 MB	2018-10-29	58b3d796d8cf96fb8580c62f46ab64d4
RStudio 1.1.463 - Mac OS X 10.6+ (64-bit)	74.5 MB	2018-10-29	a79032ba4d7daaa86a8da01948278d94
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (32-bit)	89.3 MB	2018-10-29	8a6755fa9fae2bafce289df3358aaf63
RStudio 1.1.463 - Ubuntu 12.04-15.10/Debian 8 (64-bit)	97.4 MB	2018-10-29	bc50d6bd34926c1cc3ae4a209d67d649
RStudio 1.1.463 - Ubuntu 16.04+/Debian 9+ (64-bit)	65 MB	2018-10-29	cf6d59db18619cc78d1592fefaa7c753
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (32-bit)	88.1 MB	2018-10-29	742f0bad60dfeaa3281576e14ad6699e
RStudio 1.1.463 - Fedora 19+/RedHat 7+/openSUSE 13.1+ (64-bit)	90.6 MB	2018-10-29	c7303067a0ca99deea7e427b856952d1

### Zip/Tarballs

Zip/tar archives	Size	Date	MD5
RStudio 1.1.463 - Windows Vista/7/8/10	122.9 MB	2018-10-29	1eb1d7758bd4bf4bb68d4a7c3fe8d894

Try RStudio Server Pro for free!

Fim! Escolha o sistema operacional



# Checando a Instalação Existente e os Requisitos

```
> sessionInfo()
```

```
R version 3.6.0 (2019-04-26)
```

```
Platform: x86_64-pc-linux-gnu (64-bit)
```

```
Running under: Ubuntu 18.04.2 LTS
```

```
Matrix products: default
```

```
BLAS: /usr/lib/x86_64-linux-gnu/openblas/libblas.so.3
```

```
LAPACK: /usr/lib/x86_64-linux-gnu/libopenblas-p0.2.20.so
```

```
locale:
```

```
[1] LC_CTYPE=pt_BR.UTF-8
```

```
LC_NUMERIC=C
```

```
[3] LC_TIME=pt_BR.UTF-8
```

```
LC_COLLATE=en_US.UTF-8
```

```
[5] LC_MONETARY=pt_BR.UTF-8
```

```
LC_MESSAGES=en_US.UTF-8
```

```
[7] LC_PAPER=pt_BR.UTF-8
```

```
LC_NAME=C
```

```
[9] LC_ADDRESS=C
```

```
LC_TELEPHONE=C
```

```
[11] LC_MEASUREMENT=pt_BR.UTF-8 LC_IDENTIFICATION=C
```

```
attached base packages:
```

```
[1] stats      graphics  grDevices  utils      datasets  methods   base
```

```
other attached packages:
```

```
[1] knitr_1.20
```

```
loaded via a namespace (and not attached):
```

```
[1] compiler_3.6.0 tools_3.6.0
```

## Uma instalação em Linux





# Pacotes em R



## Contributed Packages

### Available Packages

Currently, the CRAN package repository features 13853 available packages.

[Table of available packages, sorted by date of publication](#)

[Table of available packages, sorted by name](#)

### Installation of Packages

Please type `help("INSTALL")` or `help("install.packages")` in R for information on how to install packages from this repository. The manual [R Installation and Administration](#) (also contained in the R base sources) explains the process in detail.

[CRAN Task Views](#) allow you to browse packages by topic and provide tools to automatically install all packages for special areas of interest. Currently, 39 views are available.

### Package Check Results

All packages are tested regularly on machines running [Debian GNU/Linux](#), [Fedora](#), OS X, Solaris and Windows.

The results are summarized in the [check summary](#) (some [timings](#) are also available). Additional details for Windows checking and building can be found in the [Windows check summary](#).

### Writing Your Own Packages

The manual [Writing R Extensions](#) (also contained in the R base sources) explains how to write new packages and how to contribute them to CRAN.

### Repository Policies

The manual [CRAN Repository Policy \[PDF\]](#) describes the policies in place for the CRAN package repository.



### Related Directories

[Archive](#)

Previous versions of the packages listed above, and other packages formerly available.

# Mais informações sobre pacotes

Mais informações podem ser encontradas abaixo:

- <https://blog.revolutionanalytics.com/2017/01/cran-10000.html>
- [https://cran.r-project.org/web/packages/available\\_packages\\_by\\_name.html](https://cran.r-project.org/web/packages/available_packages_by_name.html)
- <https://cran.r-project.org/web/packages/>



# Baixando pacotes

```
> install.packages("ggplot2")  
> library(ggplot2)
```

Acima está o exemplo de instalação do pacote *ggplot2* utilizando o R na linha de comando. O mesmo pacote, instalado via RStudio está no exemplo abaixo.



# Instalando um Pacote via RStudio

The screenshot shows the RStudio interface with the 'Install Packages' dialog box open. The 'Repository (CRAN)' is selected, and 'ggplot2' is entered in the 'Packages' field. The 'Install dependencies' checkbox is checked. The 'Console' pane shows the R script being executed, which includes loading the 'knitr' and 'stargazer' packages, and the 'ggplot2' package. The 'Environment' pane shows the 'Global Environment' with an empty environment. The 'Packages' pane shows the 'System Library' with a list of installed packages.

**Install Packages**

Install from:  [? Configuring Repositories](#)

Packages (separate multiple with space or comma):

☒ Install dependencies

**Console**

```
> library("abind", lib.loc="/usr/local/lib/R/site-library")
> detach("package:abind", unload=TRUE)
> library("Amelia", lib.loc="/usr/local/lib/R/site-library")
Loading required package: Rcpp
##
## Amelia II: Multiple Imputation
## (Version 1.7.5, built: 2018-05-07)
## Copyright (C) 2005-2019 James Honaker, Gary King and Matthew Blackwell
## Refer to http://gking.harvard.edu/amelia/ for more information
##
> detach("package:Amelia", unload=TRUE)
>
```

**Environment**

Global Environment -

Environment is empty

**Packages**

Name	Description	Version
<b>System Library</b>		
<input type="checkbox"/> abind	Combine Multidimensional Arrays	1.4-5
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> actuar	Actuarial Functions and Heavy Tailed Distributions	2.3-1
<input type="checkbox"/> AER	Applied Econometrics with R	1.2-5
<input type="checkbox"/> Amelia	A Program for Missing Data	1.7.5
<input type="checkbox"/> aod	Analysis of Overdispersed Data	1.3
<input type="checkbox"/> archive	Multi-Format Archive and Compression Support for R Connections	1.0.0
<input type="checkbox"/> assertthat	Easy Pre and Post Assertions	0.2.0
<input type="checkbox"/> backports	Reimplementations of Functions Introduced Since R-3.0.0	1.1.2
<input type="checkbox"/> base64enc	Tools for base64 encoding	0.1-3
<input type="checkbox"/> BH	Boost C++ Header Files	1.66.0-1
<input type="checkbox"/> bindr	Parameterized Active Bindings	0.1.1
<input type="checkbox"/> bindrcpp	An 'Rcpp' Interface to Active Bindings	0.2.2
<input type="checkbox"/> bit	A Class for Vectors of 1-Bit Booleans	1.1-14
<input type="checkbox"/> bit64	A 53 Class for Vectors of 64bit Integers	0.9-7
<input type="checkbox"/> bitops	Bitwise Operations	1.0-6
<input type="checkbox"/> blob	A Simple 53 Class for Representing Vectors of	1.1.1



# Instalando Pacotes de outros sites

- A instalação de pacotes em R é feita usualmente com a função *install.packages()*
- Nesta função, informamos uma *cadeia de caracteres* com o nome do pacote que queremos instalar: por exemplo - *install.packages("yaml")*
- Há diversos tipos e pacotes em R, muitos destes pacotes podem ter sido escritos em R ou C, C++ ou Fortran.
- No caso de pacotes escritos em C, C++ e Fortran instalados com o tipo *fonte*, é necessária a compilação deste pacote.
- O R faz isso automaticamente, não é necessária, em geral nenhuma intervenção do usuário.



# Eventuais erros de compilação

```
> install.packages('fftw')
Installing package into '/usr/local/lib/R/site-library'
(as 'lib' is unspecified)
trying URL 'https://cloud.r-project.org/src/contrib/fftw_1.0-5.tar.gz'
Content type 'application/x-gzip' length 38684 bytes (37 KB)
=====
downloaded 37 KB

* installing *source* package 'fftw' ...
** package 'fftw' successfully unpacked and MD5 sums checked
checking for gcc... gcc -std=gnu99
checking whether the C compiler works... yes

...
etc
...

checking for pkg-config... /usr/bin/pkg-config
checking pkg-config is at least version 0.9.0... yes
checking for FFTW... configure: error: Package requirements (fftw3) were not met:

No package 'fftw3' found

...
etc
...

ERROR: configuration failed for package 'fftw'
* removing '/usr/local/lib/R/site-library/fftw'

The downloaded source packages are in
'/tmp/RtmpT62d7h/downloaded_packages'
Warning message:
In install.packages("fftw") :
  installation of package 'fftw' had non-zero exit status
```



Falta uma biblioteca (*fftw3*) que deve ser instalada no sistema operacional, antes de se instalar o pacote R

# Utilizando os pacotes no seu programa R

- Para utilizar um pacote em R você necessita simplesmente executar o comando *library(nomeDoPacote)*.
- Eventualmente, se você precisar retirar o pacote da memória do seu computador, se não quiser mais utilizá-lo durante a execução daquele programa é só utilizar o comando *detach*.  
Ex.: *detach(package:nomeDoPacote)*



# Mais informações sobre pacotes

- <https://www.r-bloggers.com/installing-r-packages/>
- <https://www.r-bloggers.com/how-to-install-and-include-an-r-package/>
- [http://kbroman.org/pkg\\_primer/pages/build.html](http://kbroman.org/pkg_primer/pages/build.html)
- <https://github.com/trending/r>
- <https://github.com/hadley>





# Exercícios

## Exercícios



# Dúvidas?????

Esclarecendo as Dúvidas



# Agradecimentos

Obrigado!



# Referências I



W N Venables and D M Smith.

*An Introduction to R*, volume 2.

CRAN - Comprehensive R Archive Network, 2011.

