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

Banpará

Sumário

Banpará

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/`

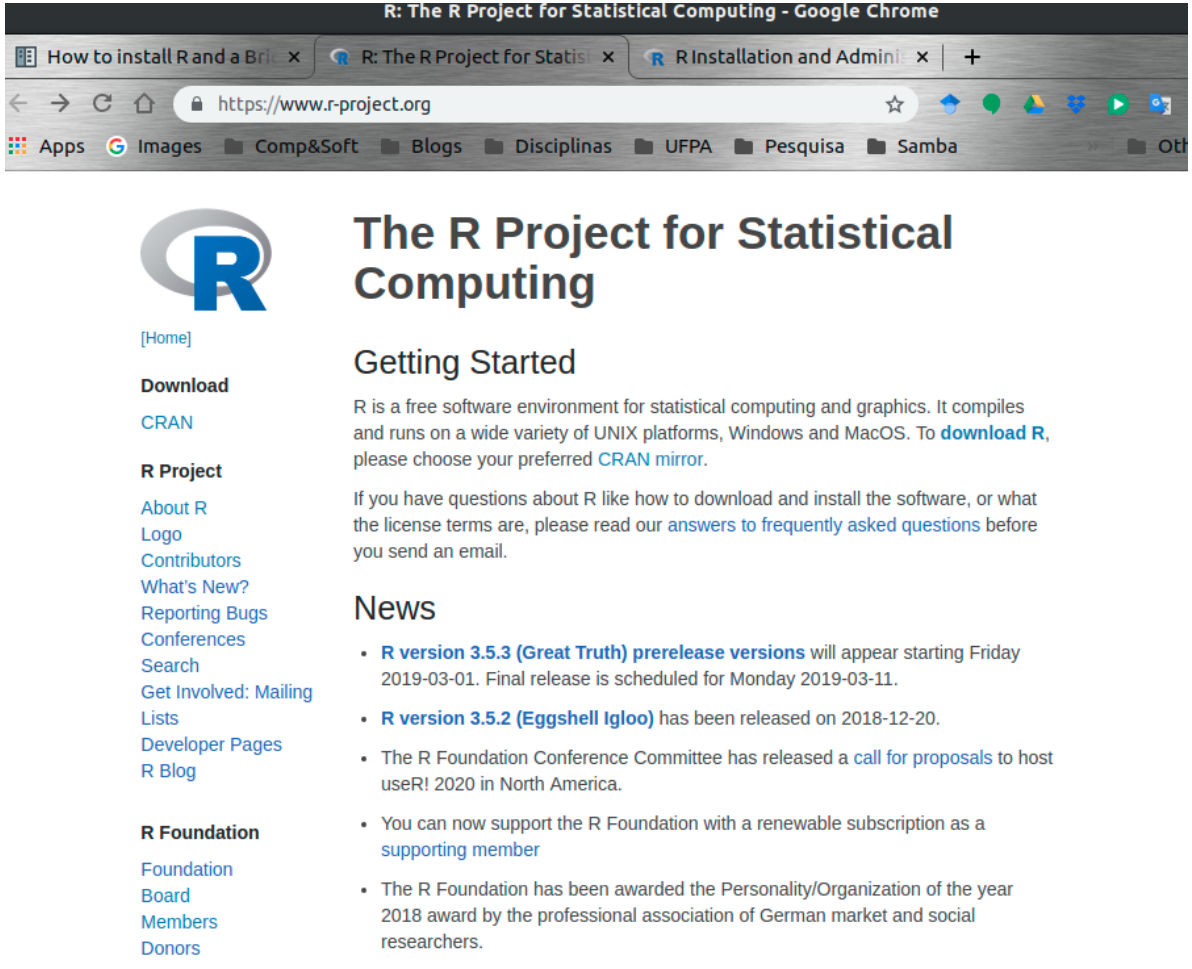
Banpará

Funcionalidades do R

Algumas das funcionalidade do R são[VS11]:

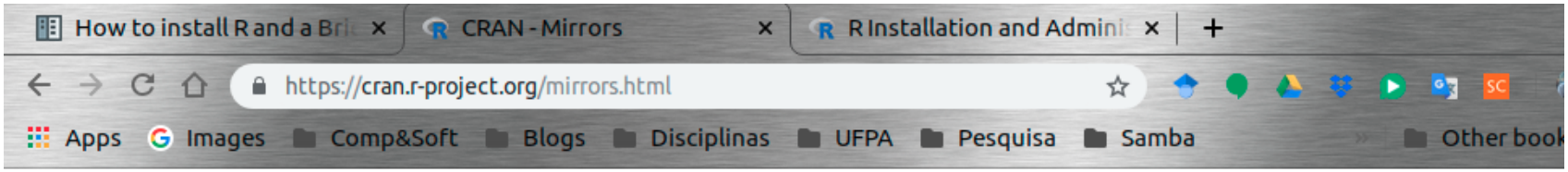
- Ferramentas para manuseio e armazenamento de dados
- Um conjunto de operadores que permitem o cálculo numérico e a manipulação de matrizes
- Um enorme conjunto de bibliotecas para análise de dados
- Ferramentas para apresentação gráfica de dados e resultados
- Uma linguagem de programação orientada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

Instalando o R: selecione **download R**



Nesta página clicar em **download R**

Instalando o R: selecione o espelho



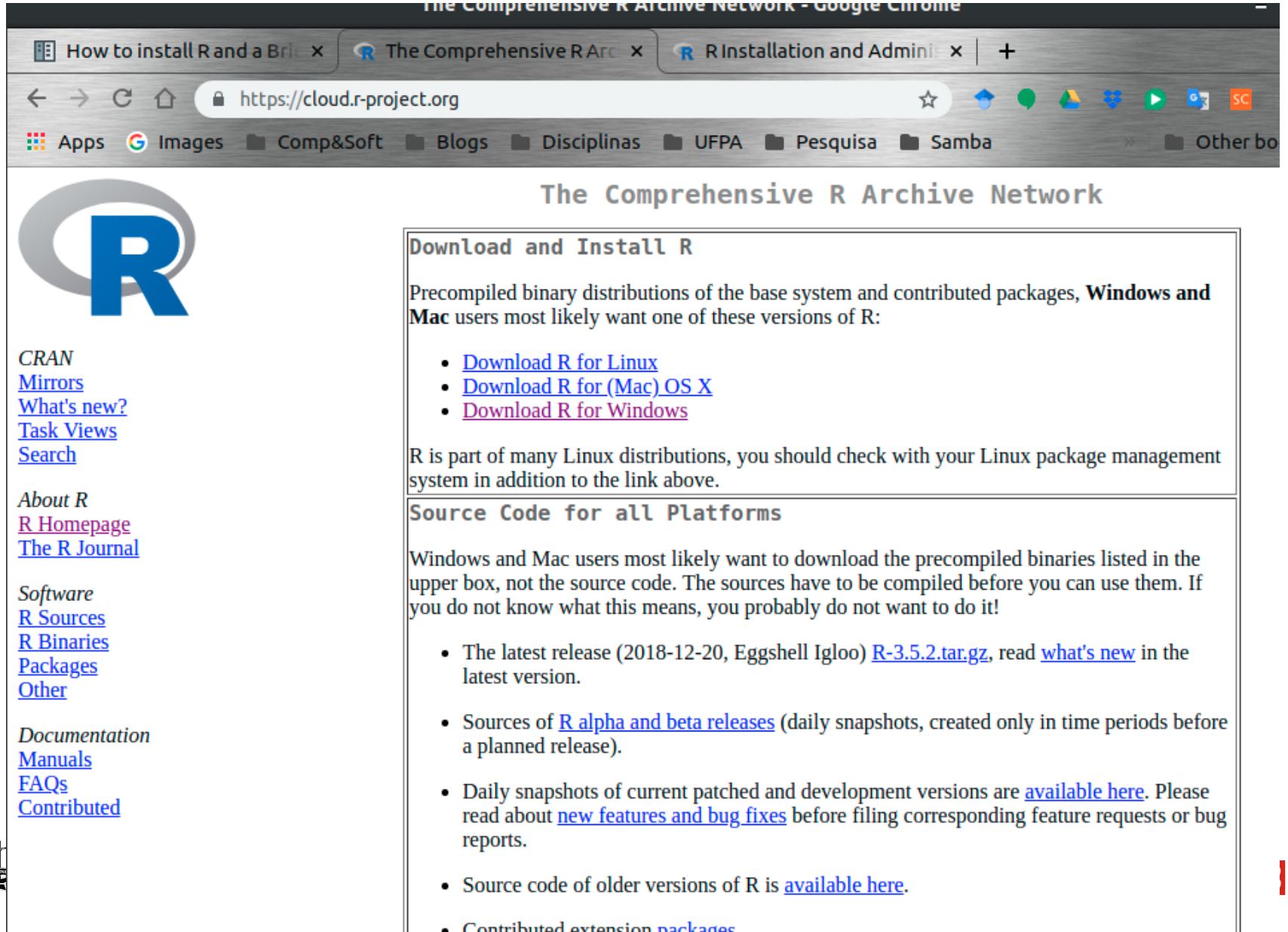
CRAN Mirrors

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.

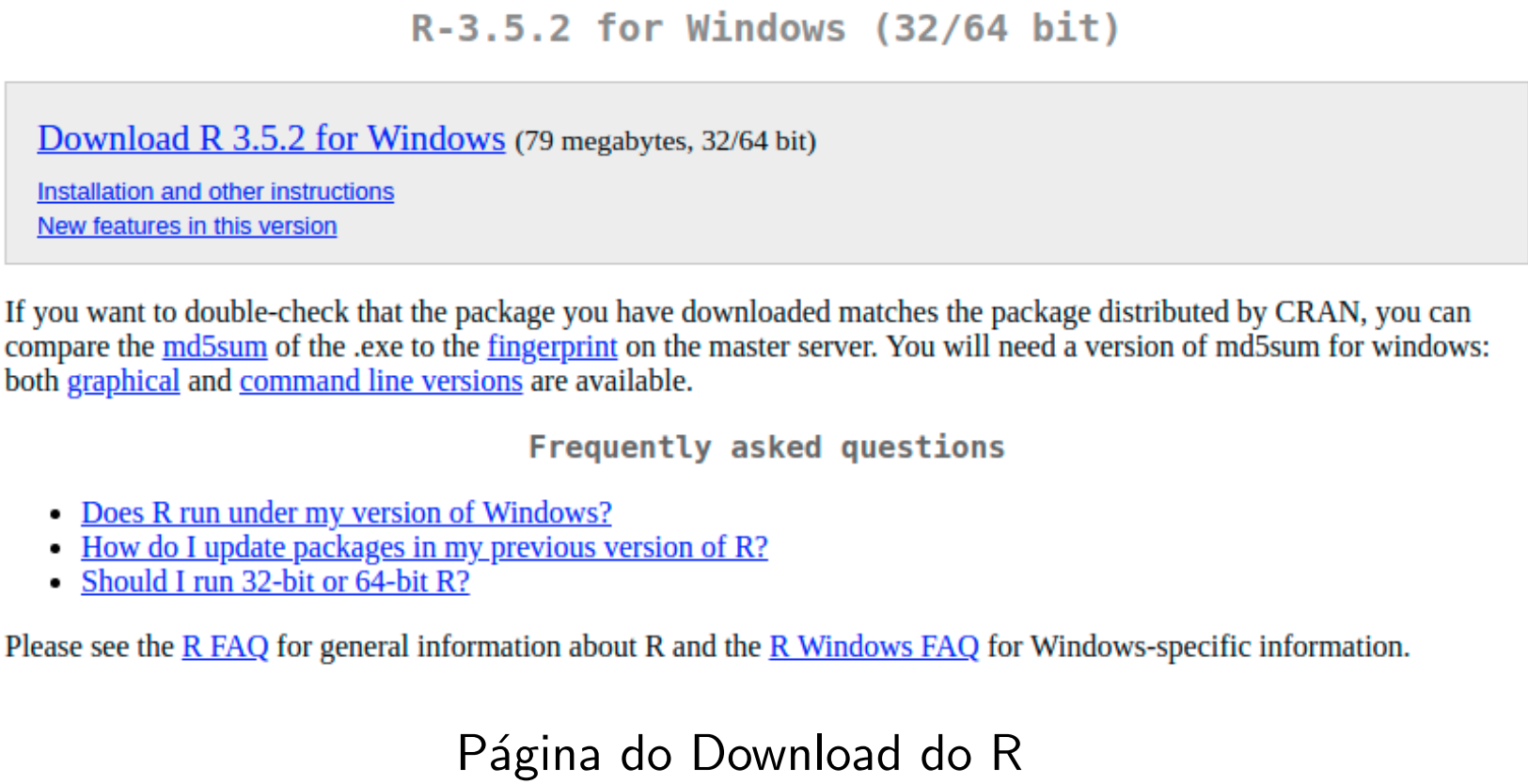
If you want to host a new mirror at your institution, please have a look at the CRAN Mirror HOWTO.

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
https://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.freestatistics.org/cran/	Patrick Wessa
http://www.freestatistics.org/cran/	Patrick Wessa

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



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

Banpará

Mais Informações sobre a Instalação do R

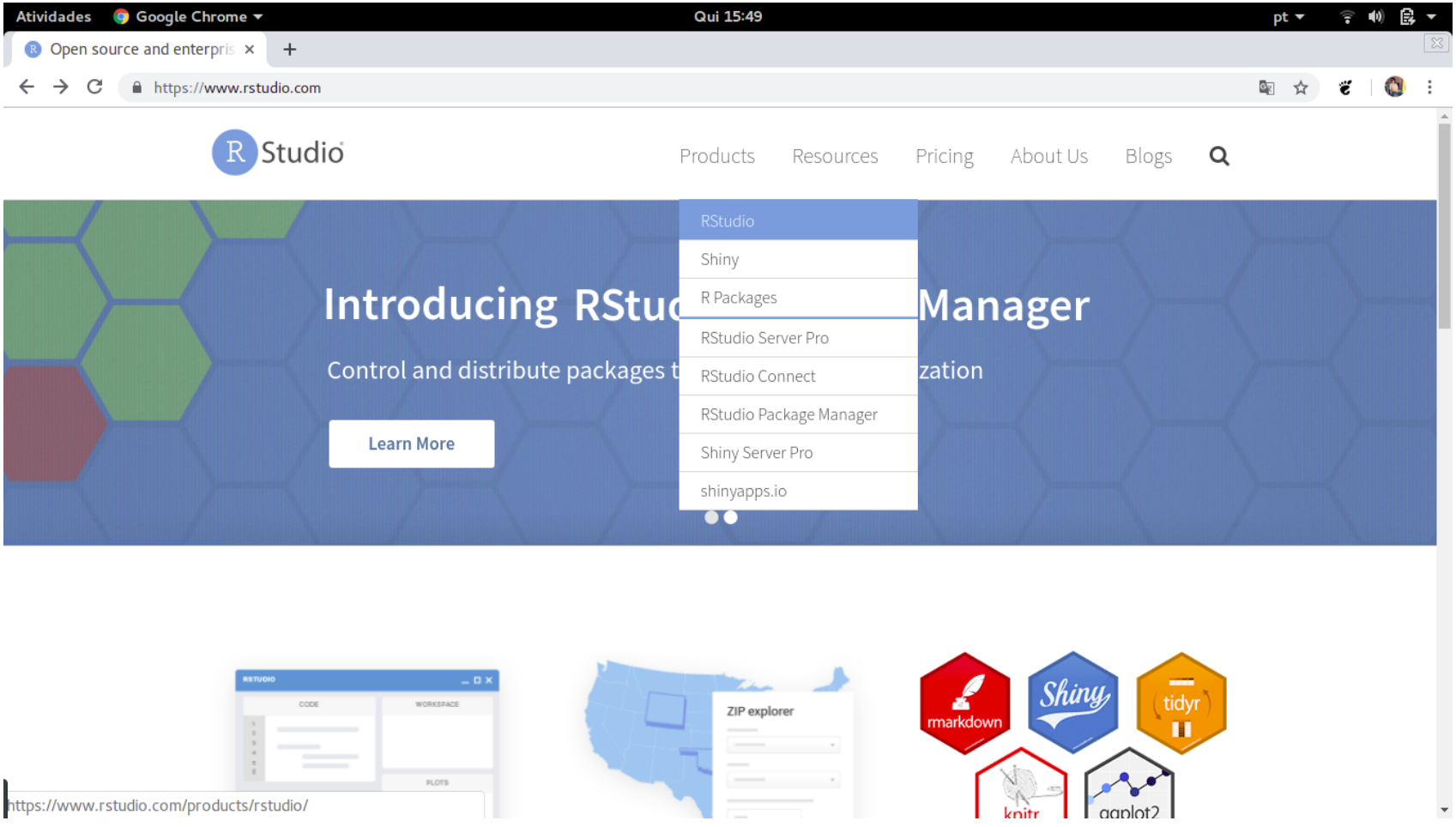
Cheque os *FAQs - Frequently Asked Questions*

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

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

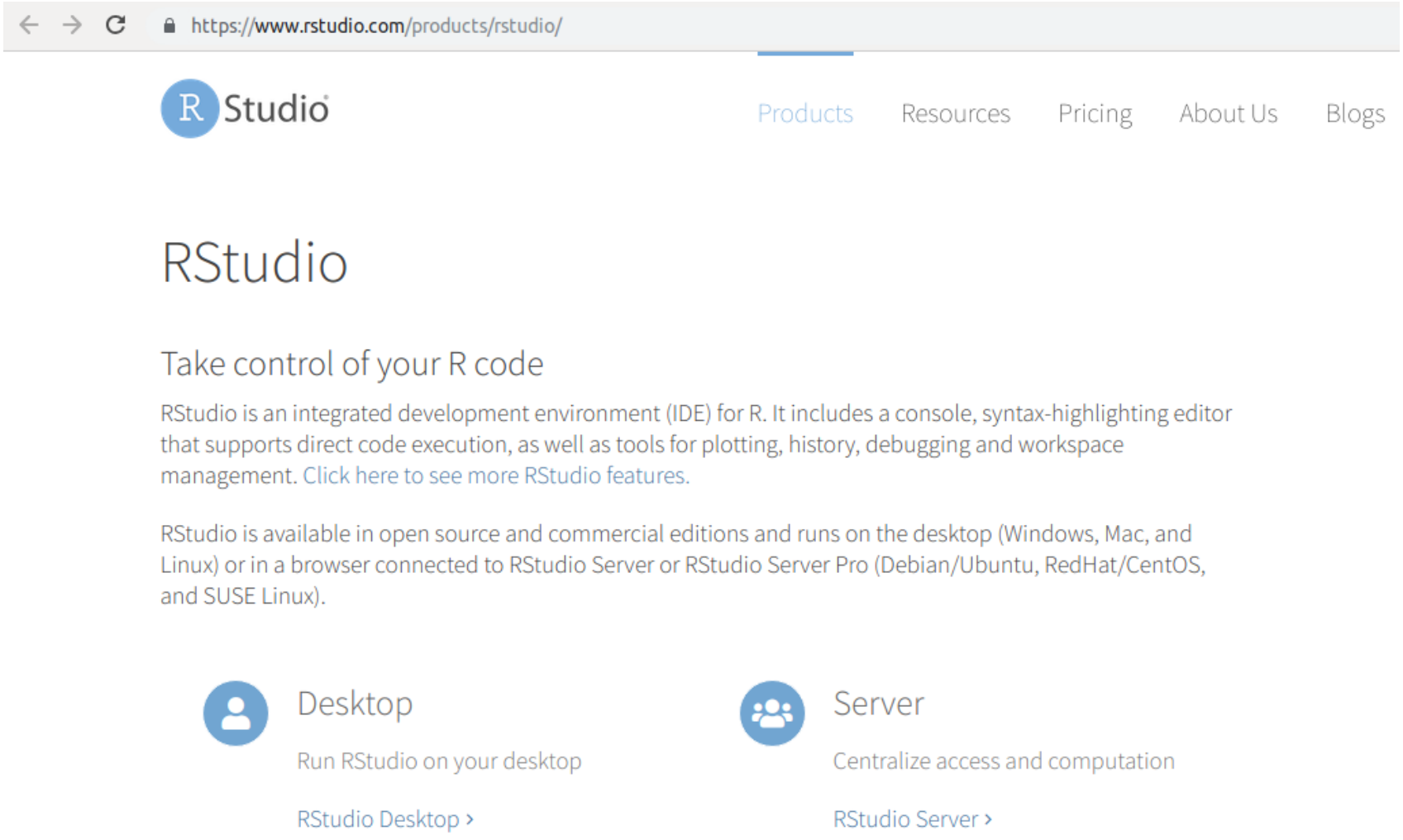
Banpará

Instalando o RStudio



https://www.rstudio.com/

Aba de Produtos do RStudio



Selecione seu RStudio

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"

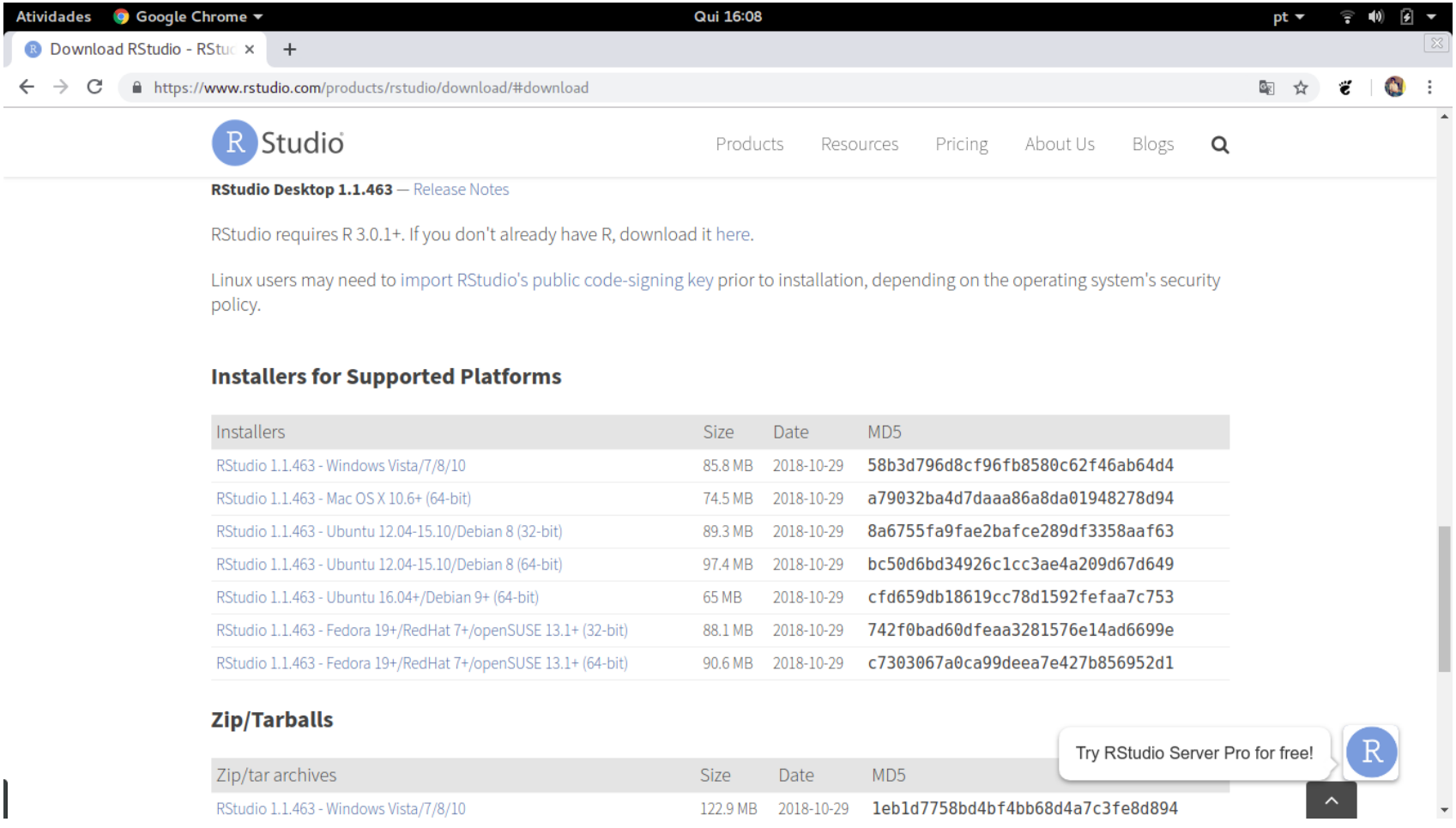
	RStudio Desktop Open Source License	RStudio Desktop Commercial License	RStudio Server Open Source License	RStudio Server Pro Commercial License	RStudio Server Pro + RStudio Connect Commercial License
	FREE	$995 per year	FREE	$9,995 per year	$29,995 per year
	DOWNLOAD	BUY	DOWNLOAD	DOWNLOAD	TALK
	Learn More	Learn More	Learn More	Learn More	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

Banpará

Finalmente, instalando seu executável



Fim! Escolha o sistema operacional

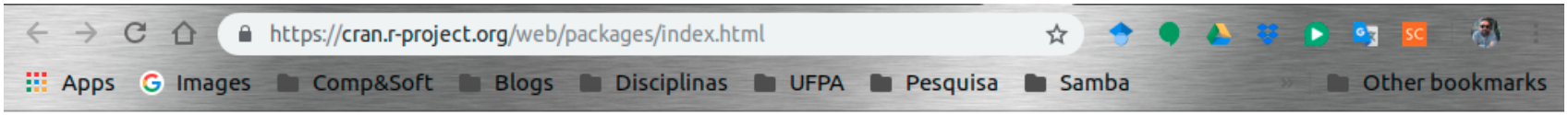
Checando a Instalação Existente e os Requisitos

```
> 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/libopenblasp-r0.2.20.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

Banpará

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.

Banpará

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/`

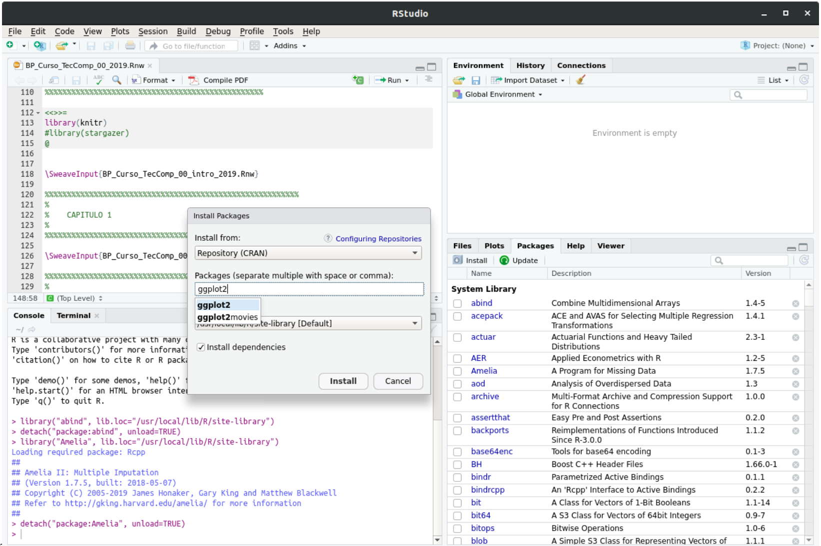
Baixando pacotes

```
> 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.

Banpará

Instalando um Pacote via RStudio



Instalando o pacote ggplot2 via Rstudio

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

Banpará

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)*

Banpará

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
- https://github.com/trending/r
- https://github.com/hadley

Banpará

Exercícios

Exercícios

Banpará

Dúvidas?????

Esclarecendo as Dúvidas

Agradecimentos

Obrigado!

Banpará

Referências I


W N Venables and D M Smith.
An Introduction to R, volume 2.
CRAN - Comprehensive R Archive Network, 2011.