

Conheça o R: Introdução, dicas e curiosidades

Gabriel Danilo Shimizu

2021-03-09

Contents

1 Conheça o R: Introdução, dicas e curiosidades	5
2 Apresentação	7
3 O que é o R	9
3.1 Porque utilizar o R?	10
4 Instalação do software R	11
4.1 Instalando RStudio	13
4.2 Primeiros passos	14
4.3 Instalando packages	16
4.4 Chamando pacote no Rstudio	18
5 Importação de dados	19
5.1 Importação de dados do excel	19
5.2 Importação de dados em csv	21
5.3 Arquivo separado por ponto e vírgula e decimal separado por vírgula	23
6 Tabulação de dados	25
6.1 DIC unifatorial	26

7 Atalhos**27**

The graphic is a promotional material for a workshop titled "Conheça o R: Introdução, dicas e curiosidades". It features a large blue "R" logo. The text "Conheça o R: Introdução, dicas e curiosidades" is displayed in white. A blue box contains the following information: "Horário e data: 19:00h - 16/03/2021", "Investimento: gratuito, entretanto sem emissão de certificado.", "INSCREVA-SE: <https://bitly.com/01BTR>", "VAGAS LIMITADAS", and "Inscrições realizadas apenas com e-mail institucional UEL". To the right, there are three diamond-shaped boxes: one for the organizer ("Msc. Gabriel Danilo Shimizu"), one for the duration ("Carga horária 3-4 horas"), and one for the number of slots ("Número de vagas: 50"). The bottom right corner features the UEL logo (a stylized building with a green dome). A circular icon with an "i" is located in the top left corner.

Conheça o R: Introdução, dicas e curiosidades

Horário e data: 19:00h - 16/03/2021
Investimento: gratuito, entretanto sem emissão de certificado.
INSCREVA-SE: <https://bitly.com/01BTR>
VAGAS LIMITADAS
Inscrições realizadas apenas com e-mail institucional UEL

Msc. Gabriel Danilo Shimizu

Carga horária 3-4 horas

Número de vagas: 50

Local: google meet

Chapter 1

Conheça o R: Introdução, dicas e curiosidades

Horário e data: 19:00h - 16/03/2021

Local: remoto via google meet

Carga horária: de 3 a 4 horas

Número de vagas: 50

Objetivo: proporcionar um conhecimento mínimo sobre a linguagem R.

Público alvo: estudantes de pós-graduação em agronomia.

Conteúdo:

- Introdução a linguagem R e ao ambiente Rstudio
- Dicas de importação de dados, geração de gráficos, atalhos, pacotes e informações no geral
- Curiosidades

Investimento: gratuito, entretanto sem emissão de certificado.

Ministrante: Gabriel Danilo Shimizu

VAGAS LIMITADAS: Inscrições realizadas apenas com e-mail institucional UEL

6CHAPTER 1. CONHEÇA O R: INTRODUÇÃO, DICAS E CURIOSIDADES

Chapter 2

Apresentação

O R é uma linguagem de programação muito utilizada na âmbito da estatística e na ciência de dados. Na Agronomia, o conhecimento sobre essa linguagem é um diferencial, sobretudo na carreira acadêmica, pois sua limitação gráfica e de análises é praticamente inexistente.

Este curso, embora curto, representa um primeiro passo para os futuros usuários de R e, é essencial para a compreensão das diversas funcionalidades dessa linguagem de programação.

Chapter 3

O que é o R

R é um ambiente computacional e uma linguagem de programação especializada em manipulação, análise e visualização gráfica de dados. Na atualidade é considerado o melhor ambiente computacional para essa finalidade. O ambiente está disponível para diferentes sistemas operacionais: Unix/Linux, Mac e Windows.

Foi criado originalmente por Ross Ihaka e por Robert Gentleman no departamento de Estatística da Universidade de Auckland, Nova Zelândia. Posteriormente, foi desenvolvido pelo esforço colaborativo de pessoas em vários locais do mundo.

O nome R provém em parte das iniciais dos criadores (Ross Ihaka e Robert Gentleman) e também de um jogo figurado com a linguagem S (da Bell Laboratories, antiga AT&T).

R é um ambiente e uma linguagem de programação similar ao S, contudo, é uma implementação distinta do S. Muitos códigos escritos para o S podem ser executados inalterados no R e vice-versa.

R é altamente expansível com o uso dos pacotes. Os pacotes são bibliotecas com dados e funções para diferentes áreas do conhecimento relacionado a estatística e áreas afins.

Um conjunto básico de pacotes vem embutido na instalação do R, com muito outros disponíveis na rede de distribuição do R (em inglês CRAN).

A linguagem R é largamente usada entre estatísticos e analistas de dados para desenvolver software de estatística e análise de dados. Pesquisas e levantamentos com profissionais da área mostram que a popularidade do R aumentou substancialmente nos últimos anos.

3.1 Porque utilizar o R?

- Software gratuito com código aberto com uma linguagem acessível; Expanção exponencial entre pesquisadores, engenheiros e estatísticos;
- Se reinventa constantemente através de novas aplicações (aproximadamente 14.762 pacotes);
- Cobertura inigualável, tecnologia de ponta;
- Totalmente flexível, permitindo desenvolver facilmente funções e pacotes para facilitar o trabalho;
- Capacidade gráfica;
- Disponível para diferentes plataformas: Windows, Linux e Mac.

Total de package no CRAN: 17275

Atualizado em: 09/03/2021

(CRAN)

Chapter 4

Instalação do software R

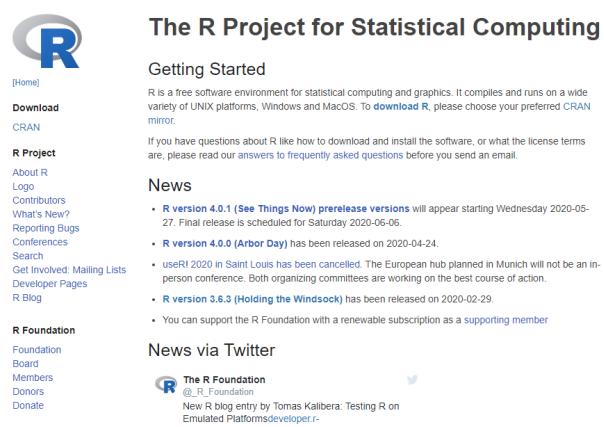
link para download:

Software R

R Studio

Acessar: <https://www.r-project.org/>

Ir em: Download > CRAN



The screenshot shows the main page of the R Project for Statistical Computing. At the top left is the R logo. To its right is the title "The R Project for Statistical Computing". Below the title is a "Getting Started" section with a brief introduction to R. Further down is a "News" section with a list of recent releases and events. At the bottom is a "News via Twitter" section showing a tweet from the official R Foundation account.

[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

Donate

... ..

The R Project for Statistical Computing

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 Mac OS. 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.

• [R version 4.0.1 \(See Things Now\) prerelease versions](#) will appear starting Wednesday 2020-05-27. Final release is scheduled for Saturday 2020-06-06.

• [R version 4.0.0 \(Arbor Day\)](#) has been released on 2020-04-24.

• userR 2020 in Saint Louis has been cancelled. The European hub planned in Munich will not be an in-person conference. Both organizing committees are working on the best course of action.

• [R version 3.6.3 \(Holding the Windsock\)](#) has been released on 2020-02-29.

• You can support the R Foundation with a renewable subscription as a [supporting member](#).

News via Twitter

 **The R Foundation** @_R_Foundation New R blog entry by Tomas Kalibera: Testing R on Emulated Platforms [developer.r](#)

Ir em: Universidade Federal do Paraná

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 .	
O-Cloud http://cloud.r-project.org/	Automatic redirection to servers worldwide, currently sponsored by Rstudio
Algeria https://cran.vsteh.dz/	University of Science and Technology Houari Boumediene
Argentina http://mirror.fcaglp.unlp.edu.ar/CRAN/	Universidad Nacional de La Plata
Austria https://cran.cnrz.ac.at/ https://mirror.sarstedt.edu.au/pub/CRAN/ https://cran.ms.unimelb.edu.au/ https://cran.curtin.edu.au/	CSIRO AARNET School of Mathematics and Statistics, University of Melbourne Curtin University
Austria https://cran.wuu.ac.at/	Wirtschaftsuniversität Wien
Belgium https://www.freestatistics.org/cran/ http://lib.ugent.be/CRAN/	Patrick Wessa Ghent University Library
Brazil https://vmbc.ufsc.br/mirror/CRAN/ https://cran.csi.ufpr.br/ https://cran.fiocruz.br/ https://vps.fluxus.usp.br/CRAN/ https://bruger.eealp.usp.br/CRAN/	Computational Biology Center at Universidade Estadual de Santa Cruz Universidade Federal do Paraná Osvaldo Cruz Foundation, Rio de Janeiro University of São Paulo, São Paulo University of São Paulo, Piracicaba
Bulgaria https://ftp.uni-sofia.bg/CRAN/	Sofia University
Canada https://mirror.cc.sfu.ca/mirror/CRAN/ https://mupr.ca/mirror/cran/ https://mirror.its.dal.ca/cran/	Simon Fraser University, Burnaby Manitoba Unix User Group Dalhousie University, Halifax

Ir em: Escolher a opção do sistema operacional do computador



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 (2020-04-24, Arbor Day) [R-4.0.0.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](#)

Questions About R

- 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.

Ir em: Instalar R pela primeira vez



R for Windows

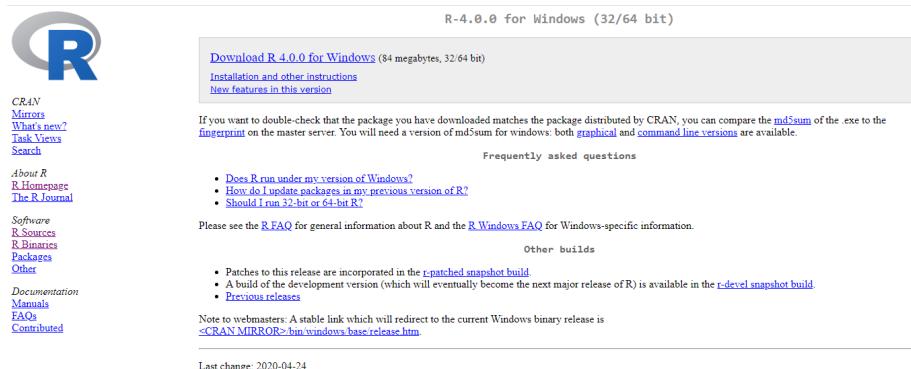
Subdirectories:

base	Binaries for base distribution. This is what you want to install R for the first time .
contrib	Binaries of contributed CRAN packages (for R >= 2.13.x; managed by Uwe Ligges). There is also information on third party software available for CRAN Windows services and corresponding environment and make variables.
old.contrib	Binaries of contributed CRAN packages for outdated versions of R (for R < 2.13.x; managed by Uwe Ligges).
Tools	Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.



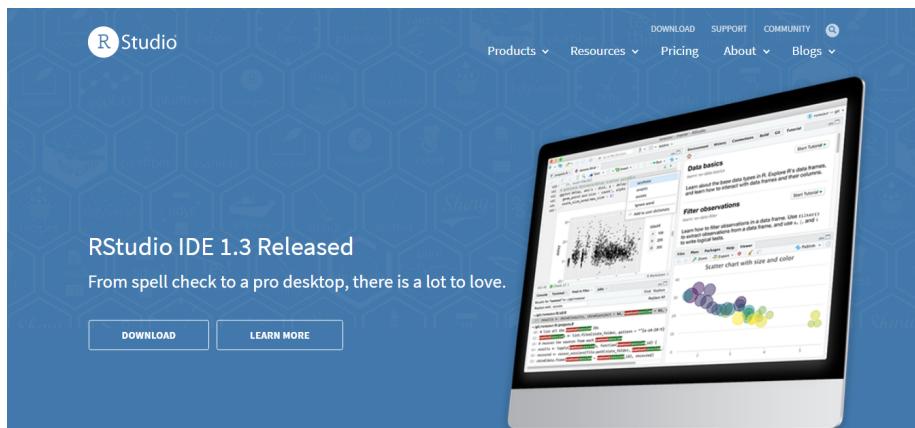
Versões anteriores do R

Executar o instalador

4.1 Instalando RStudio

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

Ir em: Download



The screenshot shows the RStudio website's download page. At the top, there's a navigation bar with links for DOWNLOAD, SUPPORT, COMMUNITY, Products, Resources, Pricing, About, and Blogs. The main heading is "Download RStudio". Below it, a section titled "Choose Your Version" explains what RStudio is and its features. To the right, there's a promotional image for "RStudio Team" featuring a stylized "R" logo and some text about RStudio's new solution for professional data science teams.

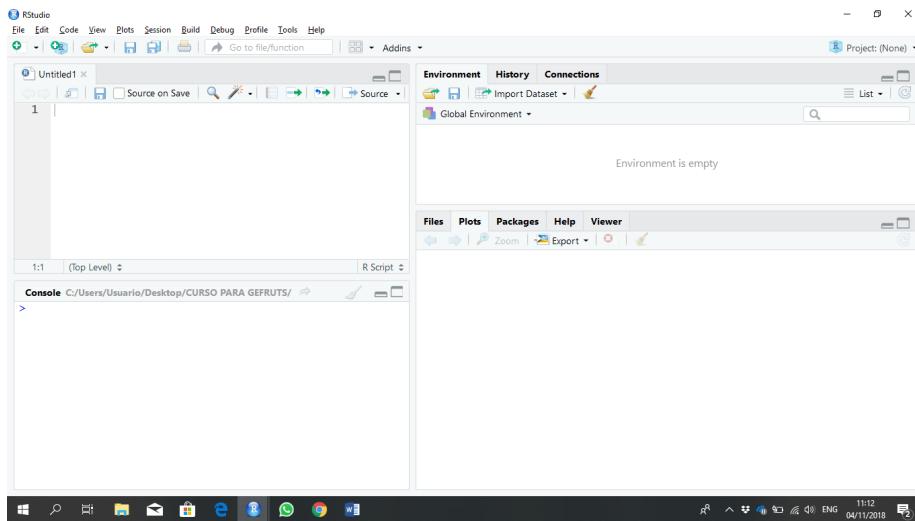
Baixar a versão do Rstudio correspondente ao seu sistema operacional

OS	Download	Size	SHA-256
Windows 10/8/7	RStudio-1.3.959.exe	171.41 MB	3d493ae5
macOS 10.13+	RStudio-1.3.959.dmg	148.57 MB	7c5b695d
Ubuntu 16	rstudio-1.3.959-amd64.deb	124.57 MB	c2931495
Ubuntu 18/Debian 10	rstudio-1.3.959-amd64.deb	126.11 MB	411ab500
Fedora 19/Red Hat 7	rstudio-1.3.959-x86_64.rpm	146.24 MB	a144e4e6
Fedora 28/Red Hat 8	rstudio-1.3.959-x86_64.rpm	150.32 MB	57169bee
Debian 9	rstudio-1.3.959-amd64.deb	126.42 MB	b2d9366f
SLES/OpenSUSE 12	rstudio-1.3.959-x86_64.rpm	119.02 MB	bcc9387e
OpenSUSE 15	rstudio-1.3.959-x86_64.rpm	127.59 MB	a4f404f0

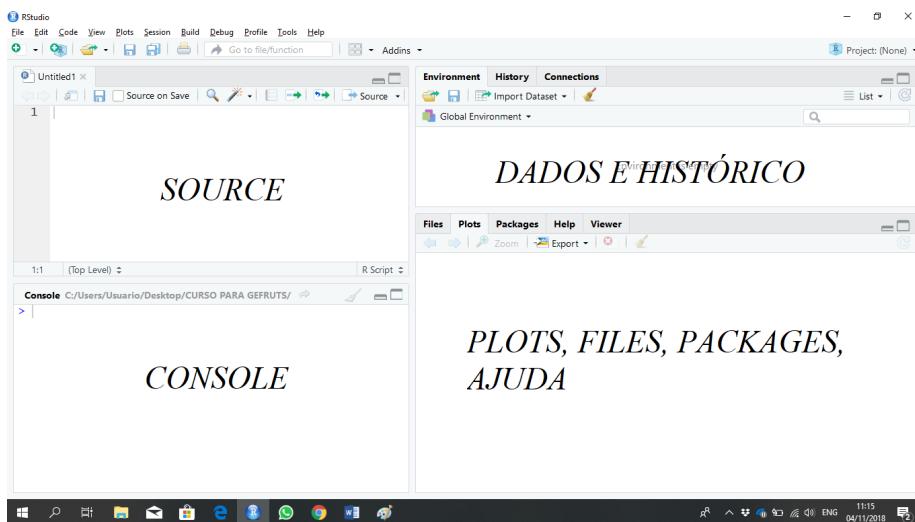
Executar o instalador.

4.2 Primeiros passos

Abra o Rstudio

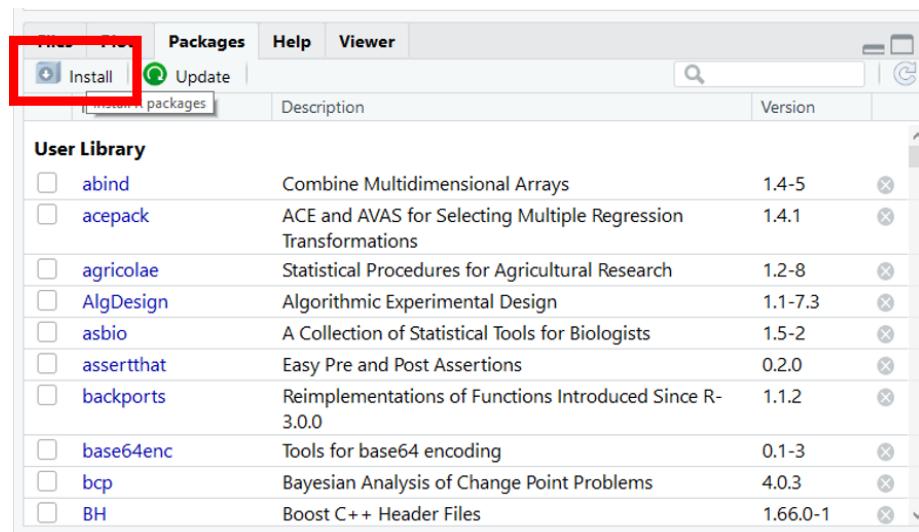
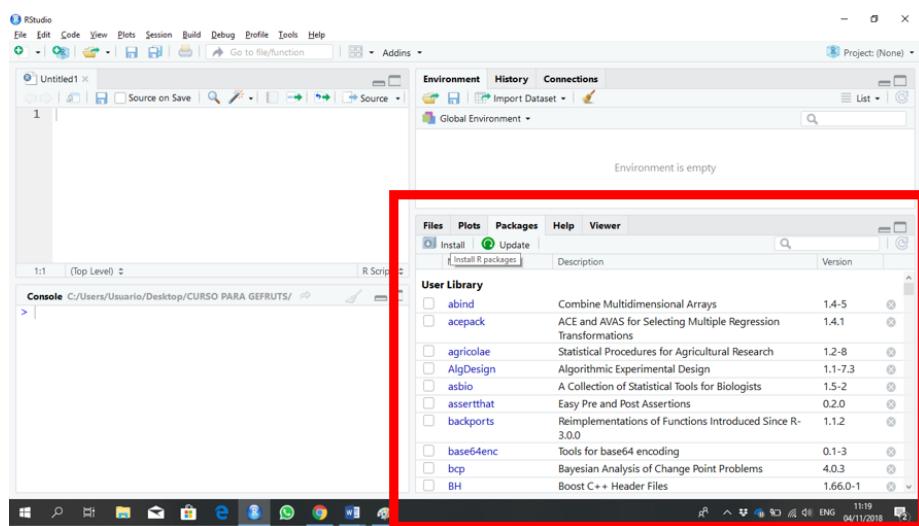


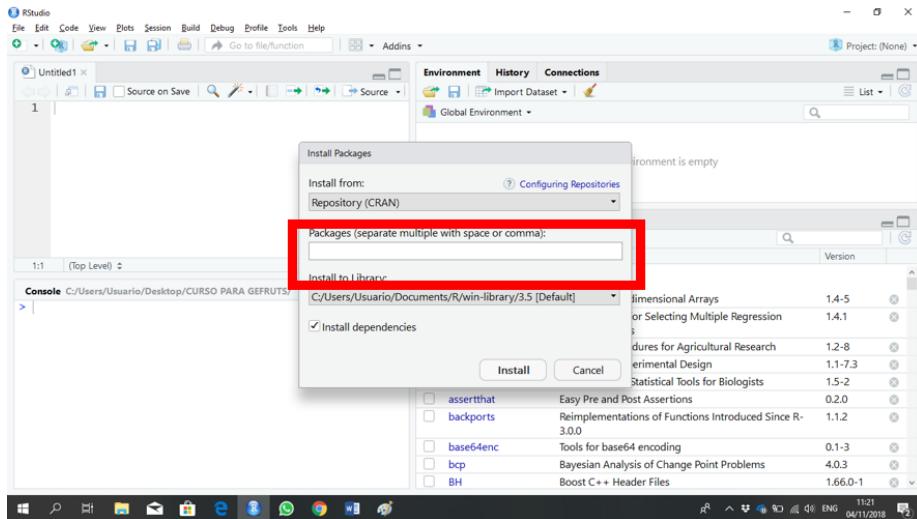
Ambiente Rstudio



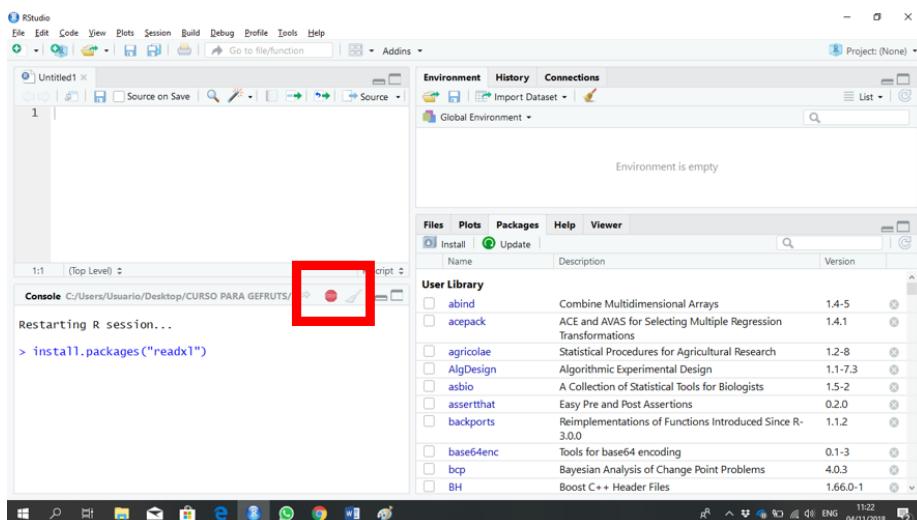
Source: é seu script (Sempre construir o script aqui, nunca no console) Console: é a saída Dados e histórico: é onde está os dados e tudo que foi realizado durante a análise Plots, files, packages, ajuda: é a saída gráfica, as pastas do diretório atual, os pacotes instalados e a ajuda

4.3 Instalando packages

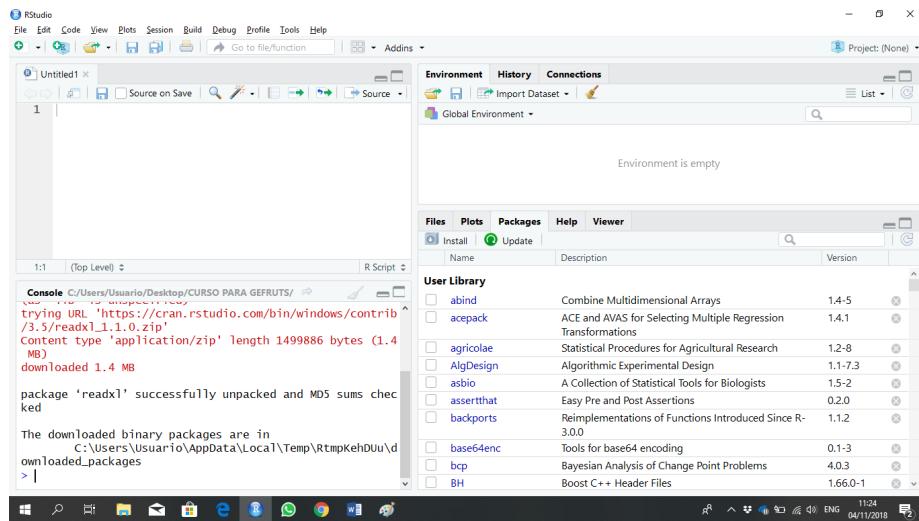




Digitar o nome do pacote desejado e depois em “Install”.



Toda vez que aparecer o ícone em vermelho, o Rstudio está trabalhando, dessa forma, não executar mais nada até o ícone desaparecer.



4.4 Chamando pacote no Rstudio

Função:

```
library(nome do pacote) require(nome do pacote)
```

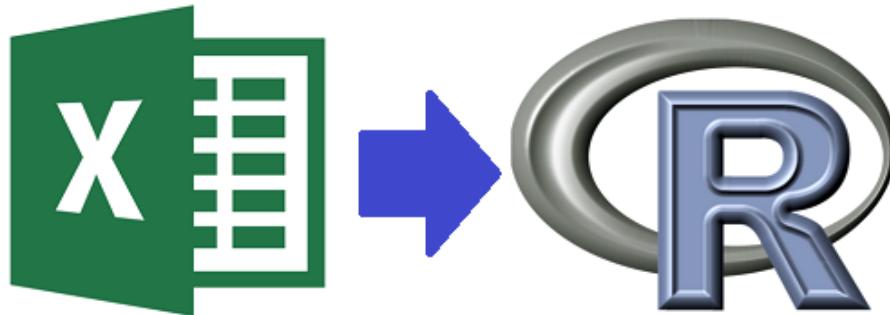
nome do pacote::

Ex. library(readxl); require(readxl); readxl::

Chapter 5

Importação de dados

5.1 Importação de dados do excel



5.1.1 Utilizando a package `readxl`

Existem diversas formas de importação arquivos para o R. Uma das mais utilizadas é a importação de um arquivo em excel.

O excel pode possuir dois tipos de extensão, `.xls` ou `.xlsx`, sendo as versões anteriores e superiores ao office 2010, respectivamente.

Uma das packages mais utilizadas para importação os dados de um arquivo em excel é chamado de “`readxl`”.

Para importar por esse pacote, devemos seguir os passos a seguir:

5.1.2 Especificar o diretório onde fica o arquivo em excel.

Uma das formas de especificar o diretório buscando a pasta manualmente, como na Figura 1 (copiar e colar no Source do Rstudio).

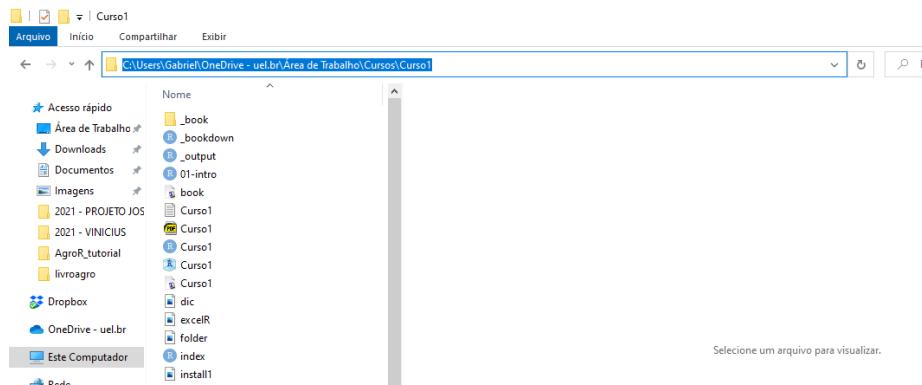


Figure 5.1: Figura 1: Especificando diretório de trabalho

A seguir, é necessário trocar as barras de \ para / ou adicionar mais uma \ (\) para \() e colocar dentro do comando setwd e o diretório entre aspas.

Obs. O diretório abaixo é do **meu computador!!!**

```
setwd("C:\\\\Users\\\\Gabriel Shimizu\\\\Dropbox\\\\ProjetoExperimental\\\\Análise Descritiva")  
## ou  
setwd("C:/Users/Gabriel Shimizu/Dropbox/ProjetoExperimental/Análise Descritiva")
```

Obs. Sempre que for alterar a local da pasta, devemos alterar a localização do diretório!!!

5.1.3 Opção 2 para especificar diretório

Atalho: ctrl+shift+h

Este comando irá buscar as **pastas** (não irá aparecer nenhum arquivo a menos que seja uma pasta). Buscar a pasta que contém o arquivo (Deverá saber onde fica, pois por esse método, o arquivo em extensão excel não irá aparecer, uma vez que o excel não é uma pasta)

Após encontrar, clicar em “open”. No console do R irá aparecer **setwd(localização)**, esta é a localização. Recomendo copiar e colar no Source do Rstudio

5.1.4 Conferir se o arquivo está no diretório especificado

```
dir()
```

5.1.5 Ativando a package e importando o arquivo

Antes de ativar o pacote, deve-se instalar o mesmo (Ver guia de instalação - Instalação)

```
library(readxl)  
dados=read_excel("DIC.xlsx", sheet=1)
```

O argumento sheet=1, está se referindo a planilha 1 do arquivo em excel (podemos exportar de outras planilhas)

Obs. No caso de extensão “.xls”, não esquecer de mudar!!!

5.1.6 Ajuda

```
?readxl
```

Conjunto de dados

5.2 Importação de dados em csv

Existem diversas formas de exportar arquivos para o R. Uma das mais utilizadas é a importação de um arquivo em extensão .csv.

Para importar dados de um arquivo em .csv, podemos usar os comandos read.csv (arquivo separado por vírgula e decimal por ponto) ou read.csv2 (Arquivo separado por ponto e vírgula e decimal por vírgula).

Para importar dados em extensão .csv, devemos seguir os passos a seguir:

5.2.1 Especificar o diretório onde fica o arquivo de dados.

Uma das formas de especificar o diretório buscando a pasta manualmente, como na Figura 1 (copiar e colar no Source do Rstudio).

A seguir, é necessário trocar as barras de / ou adicionar mais uma \ (\ para \) e colocar dentro do comando setwd e o diretório entre aspas.

Obs. O diretório abaixo é do meu computador!!!

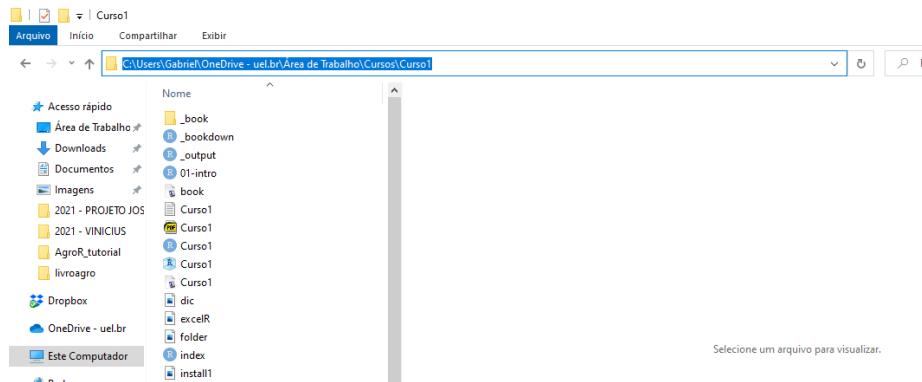


Figure 5.2: Figura 1: Especificando diretório de trabalho

```
setwd("C:\\\\Users\\\\Gabriel Shimizu\\\\Dropbox\\\\SITE\\\\EXPERIMENTAL")
## ou

setwd("C:/Users/Gabriel Shimizu/Dropbox/SITE/EXPERIMENTAL")
```

Obs. Sempre que for alterar a local da pasta, devemos alterar a localização do diretório!!!

5.2.2 Opção 2 para especificar diretório

Atalho: **ctrl+shift+h**

Este comando irá buscar as **pastas** (não irá aparecer nenhum arquivo a menos que seja uma pasta). Buscar a pasta que contém o arquivo (Deverá saber onde fica, pois por esse método, o arquivo em extensão excel não irá aparecer, uma vez que o excel não é uma pasta)

Após encontrar, clicar em “**open**”. No console do R irá aparecer **setwd(localização)**, esta é a localização. Recomendo copiar e colar no **Source** do Rstudio

5.2.3 Conferir se o arquivo está no diretório especificado

```
dir()
```

5.2.4 Importando o arquivo

5.3 Arquivo separado por ponto e vírgula e decimal separado por vírgula

```
dados=read.csv2("DIC (CSV ponto e virgula).csv")
dados
```

5.3.1 Arquivo separado por vírgula e decimal com ponto

```
dados=read.csv("DIC (CSV virgula).csv", sep=",")
dados
```

Dados (Separado por vírgula)

Dados (Separado por ponto e vírgula)

Chapter 6

Tabulação de dados

Apesar da simplicidade em se tabular dados em uma planilha excel, a grande maioria dos acadêmicos tem dificuldade em se efetuar tal tarefa. Isso torna-se ainda pior, quando os mesmos precisam tabular de uma forma específica para um determinado *Software*. Nesse contexto, o presente tutorial tem a finalidade de auxiliar os usuários de R a estruturar a planilha em excel de tal forma a facilitar as análises com ênfase em experimentação agronômica.

O que não colocar em sua planilha!

- Frequentemente é comum que os usuários de excel realizem cálculos de medidas de posição e dispersão, tais como média, variância, desvio-padrão, etc... Entretanto, essas células preenchidas por tais estatísticas de nada contribuem para quem irá trabalhar com o R, muito pelo contrário, acabam gerando mais trabalho, visto que em alguns casos podem ocasionar confundimento no *Software*;
- Deve-se evitar nomes de colunas muito extensos, pois operacionalmente digitar tais nomes pode gerar complicações futuras;
- Nome de colunas sempre na primeira linha;
- Evitar nome dos níveis do fator (Tratamentos) como numérico (1,2,3,4,...), exceto quando os tratamentos são quantitativos;
- Evitar pular células (Células em branco), a menos que tenha dados faltantes (parcelas perdidas).

Monte a planilha da forma mais simples possível!!!

6.1 DIC unifatorial

Experimentos em delineamento inteiramente casualizado só possuem o tratamento como fator. Dessa forma, em uma planilha, só necessitamos de **uma coluna de tratamentos e uma coluna de resposta**.

Quando há mais de uma variável resposta, pode-se adicionar as variáveis em cada coluna, lado a lado.

Abaixo, segue um imagem de como montar um arquivo em excel de um experimento em DIC com seis tratamentos e quatro repetições e o link para o download de um arquivo em excel (extensão .xlsx)

Tratamento	Resposta 1	Resposta 2	Resposta 3	Resposta 4	Resposta 5	Resposta 6
T1						
T1						
T1						
T1						
T2						
T2						
T2						
T2						
T3						
T3						
T3						
T3						
T4						
T4						
T4						
T4						
T5						
T5						
T5						
T5						
T6						
T6						
T6						
T6						

Chapter 7

Atalhos

Existem vários atalhos no Rstudio que facilitam a vida do usuário. Nos capítulos anteriores, vimos que o atalho ctrl+shift+h facilita a especificação do diretório de trabalho. Outros atalhos úteis, tais como o ctrl+shift+k para geração de relatórios via Rmarkdown, ou ctrl+shift+c para adicionar # sobre as linhas são bastante populares entre os usuários.

A lista completa de atalhos do Rstudio podem ser acessadas utilizando o alt shift k.

Particularmente eu utilizo comandos como os já mencionados, o ctrl+shift+n para abrir uma nova guia R; ctrl+A para selecionar tudo; ctrl+shift+home ou ctrl+shift+end para selecionar da linha até o começo ou da linha até o fim, respectivamente; ctrl+S para salvar; ctrl+x para recortar; ctrl+z, ctrl+c e ctrl+v para voltar alteração, copiar e colar, respectivamente.

Assim como mencionado, existe diversos atalhos disponíveis no Rstudio, alguns pouco funcionais.