Espectroscopia em herbário com MicroNIR: Teoria e prática

Aula 2 (tarde)

Caroline Vasconcelos

Dra. em Botânica, Bolsista AT – FAPEAM/ Herbário INPA

Manaus, 03 de abril de 2025

Conteúdo

- Reunindo as leituras obtidas por cada grupo
- Aula expositiva sobre R:
  - Boas práticas no R
  - Executando um script R:
    - Instalação das bibliotecas necessárias
    - Importação e preparação das leituras espectrais
    - Inspeção visual dos espectros
    - Preparação dos conjuntos de dados para análise
- Mãos à obra!

Reunião das leituras obtidas por cada grupo

- Os grupos devem repassar as leituras feitas

- Lembrando que as leituras estão dentro da pasta de trabalho criada por vocês em:
  - Documentos/InnoSpectra/Scan Results/

Quais arquivos vamos precisar?

- Leituras espectrais (vários arquivos **.csv**)
- Script (**01_Importando e preparando dados espectrais.R**)
- Banco de dados de *Handroanthus* do herbário INPA (**Handroanthus.xlsx**)

Boas práticas no R

- Considere instalar numa pasta segura (por exemplo, C:\R):
  - R tools - https://cran.r-project.org/bin/windows/Rtools/

cran.r-project.org/bin/windows/Rtools/

Daily use | Google | Tasks | Courses | Applications | Submissions | Reviews | Papers | Tools | Blogs | Books | Bases | HerbTools | Events | MS | Todos os favoritos

RTools: Toolchains for building R and R packages from source on Windows

Choose your version of Rtools:

RTools 4.5	for R versions from 4.5.0 (R-prerelease and R-devel)
RTools 4.4	for R versions 4.4.x (R-release)
RTools 4.3	for R versions 4.3.x (R-oldrelease)
RTools 4.2	for R versions 4.2.x
RTools 4.0	for R from version 4.0.0 to 4.1.3
old versions of RTools	for R versions prior to 4.0.0

Rtools45 for Windows

Rtools is a toolchain bundle used for building R packages from source (those that need compilation of C/C++ or Fortran code) and for building R itself. Rtools45 is currently used for R 4.5 and R-devel, the development version of R, to become R 4.6.0.

Rtools45 consists of Msys2 build tools, GCC 14/MinGW-w64 compiler toolchain, libraries built using the toolchain, and QPDF. Rtools45 supports 64-bit Windows and UCRT as the C runtime.

Compared to Rtools44, Rtools45 for 64-bit Intel machines has newer versions of two core components: GCC and binutils. It is recommended to re-compile all code with the new toolchain to avoid problems.

Rtools45 is also available for 64-bit ARM machines (aarch64): it includes Msys2 build tools (64-bit Intel builds running via emulation) and aarch64 builds of LLVM 19/MinGW-w64 compiler toolchain, libraries built using the toolchain, and again QPDF. The 64-bit ARM version of Rtools45 is experimental: a number of CRAN packages don't work with it and the Fortran compiler (flang-new) is not yet able to compile Fortran code of all CRAN packages. A number of CRAN packages don't work because they require not-yet-available 64-bit ARM versions of external software.

Installing Rtools45

Rtools is only needed for installation of R packages from source (those that need compilation of C/C++ of Fortran code) or building R from source. R can be installed from the R binary installer and by default will install binary versions of CRAN packages, which does not require Rtools45.

Moreover, online build services are available to check and build R packages for Windows, for which again one does not need to install Rtools45 locally. The Winbuilder check service uses identical setup as the CRAN incoming packages checks and has already all CRAN and Bioconductor packages pre-installed.

Rtools45 may be installed from the Rtools45 installer or 64-bit ARM Rtools45 installer. It is recommended to use the defaults, including the default installation location of `C:\rtools45`.

When using R installed by the installer, no further setup is n after installing Rtools45 to build R packages from source. When using the default installation location, R and Rtools45 may be installed in any order and Rtools45 may be installed when R is already running.

On ARM, binary versions of packages are currently not available from CRAN, so Rtools45 is required to install any package that needs compilation.

Additional information

A detailed tutorial on how to build R and packages using Rtools45 for R package authors and R developers is available for R 4.5.x and R-devel.

From the user perspective, Rtools45 is the same as Rtools42-44. However, it uses newer versions of the compiler toolchain and libraries and some libraries have been added, hence some package authors will have to adapt their make files. Maintainers of CRAN and Bioconductor packages may use these patches for reference or re-use them in their code.

A change log for Rtools45 vs Rtools44 and of individual revisions of Rtools45 is available here

Rtools45 is also available in base and full toolchain tarballs suitable for users who have their own installation of Msys2. The base toolchain tarball is smaller and includes only what is needed to build R and the recommended packages. All Rtools files are available here

Boas práticas no R

- Considere instalar numa pasta segura (por exemplo, C:\R):
  - R tools - https://cran.r-project.org/bin/windows/Rtools/
  - R base - https://cran.r-project.org/bin/windows/base/

cran.r-project.org/bin/windows/base/

Daily use | Google | Tasks | Courses | Applications | Submissions | Reviews | Papers | Tools | Blogs | Books | Bases | HerbTools | Events | MS | Todos os favoritos

R-4.4.3 for Windows

[Download R-4.4.3 for Windows](#) (85 megabytes, 64 bit)

README on the Windows binary distribution
New features in this version

This build requires UCRT, which is part of Windows since Windows 10 and Windows Server 2016. On older systems, UCRT has to be installed manually from here.

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.

Frequently asked questions

- Does R run under my version of Windows?
- How do I update packages in my previous version of R?

Please see the R FAQ for general information about R and the R Windows FAQ for Windows-specific information.

Other builds

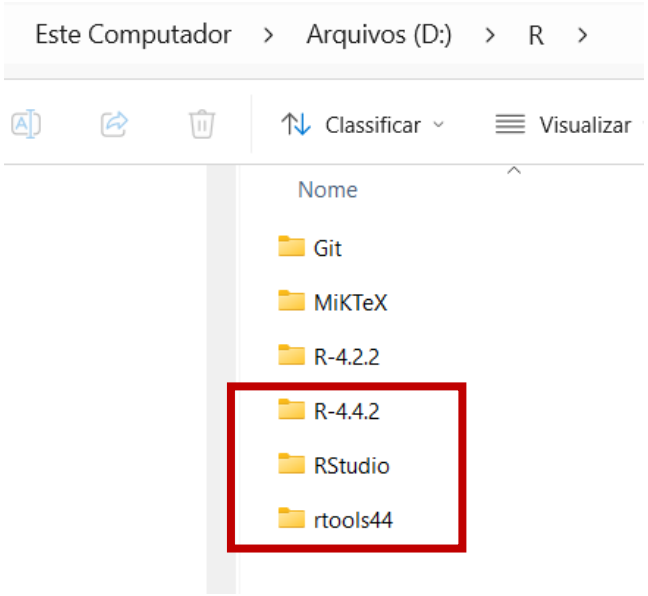
- A pre-release version for the forthcoming R-4.5.0 is available.
- Patches to this release are incorporated in the r-patched snapshot build.
- A build of the development version (which will eventually become the next major release of R) is available in the r-devel snapshot build.
- Previous releases

Note to webmasters: A stable link which will redirect to the current Windows binary release is <CRAN MIRROR>/bin/windows/base/release.html.

Last change: 2025-03-14

Boas práticas no R

- Considere instalar numa pasta segura (por exemplo, C:\R):
  - R tools - https://cran.r-project.org/bin/windows/Rtools/
  - R base - https://cran.r-project.org/bin/windows/base/
  - R Studio (IDE) - https://posit.co/download/rstudio-desktop/



DOWNLOAD

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

Don't want to download or install anything? Get started with RStudio on Posit Cloud for free. If you're a professional data scientist looking to download RStudio and also need common enterprise features, don't hesitate to book a call with us.

Want to learn about core or advanced workflows in RStudio? Explore the RStudio User Guide or the Getting Started section.

1: Install R

RStudio requires R 3.6.0+. Choose a version of R that matches your computer's operating system.

R is not a Posit product. By clicking on the link below to download and install R, you are leaving the Posit website. Posit disclaims any obligations and all liability with respect

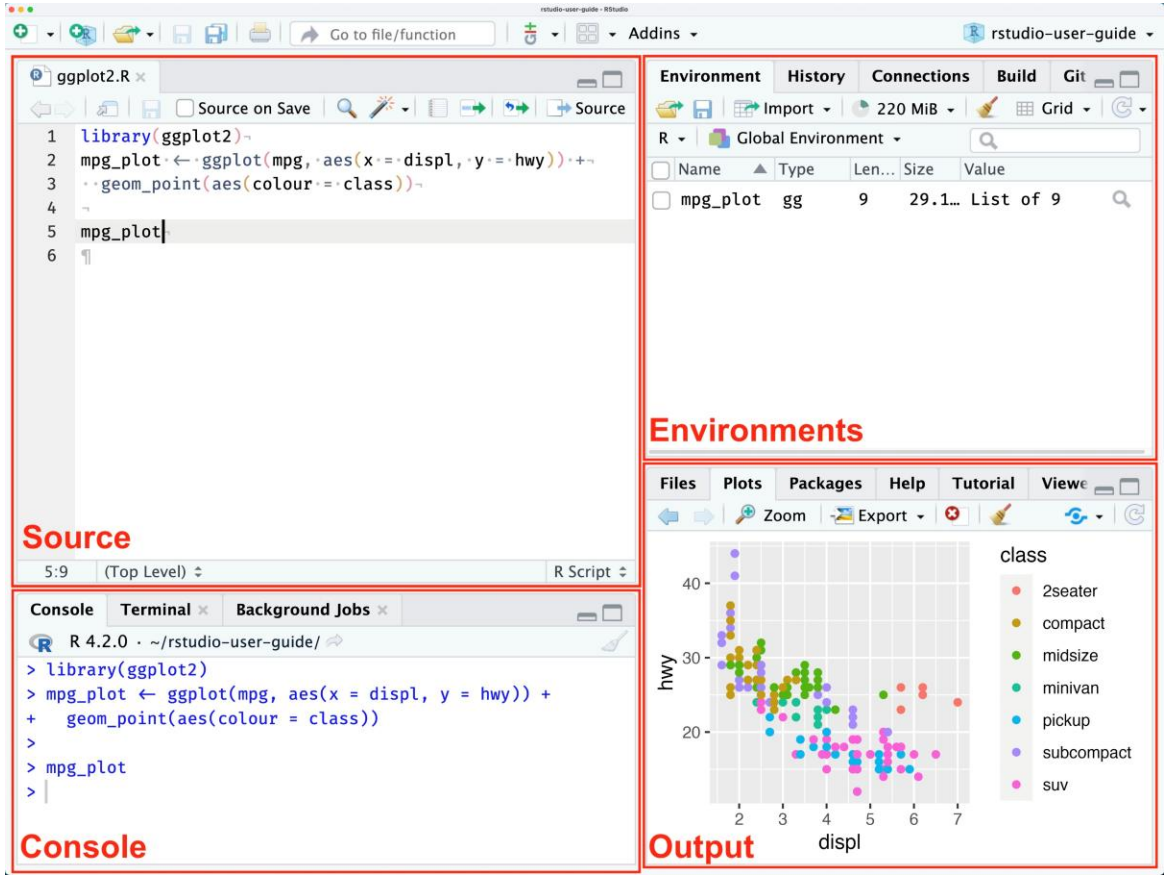
2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 265.28 MB | SHA-256: BB369743 | Version: 2024.12.1+563
Released: 2025-02-13

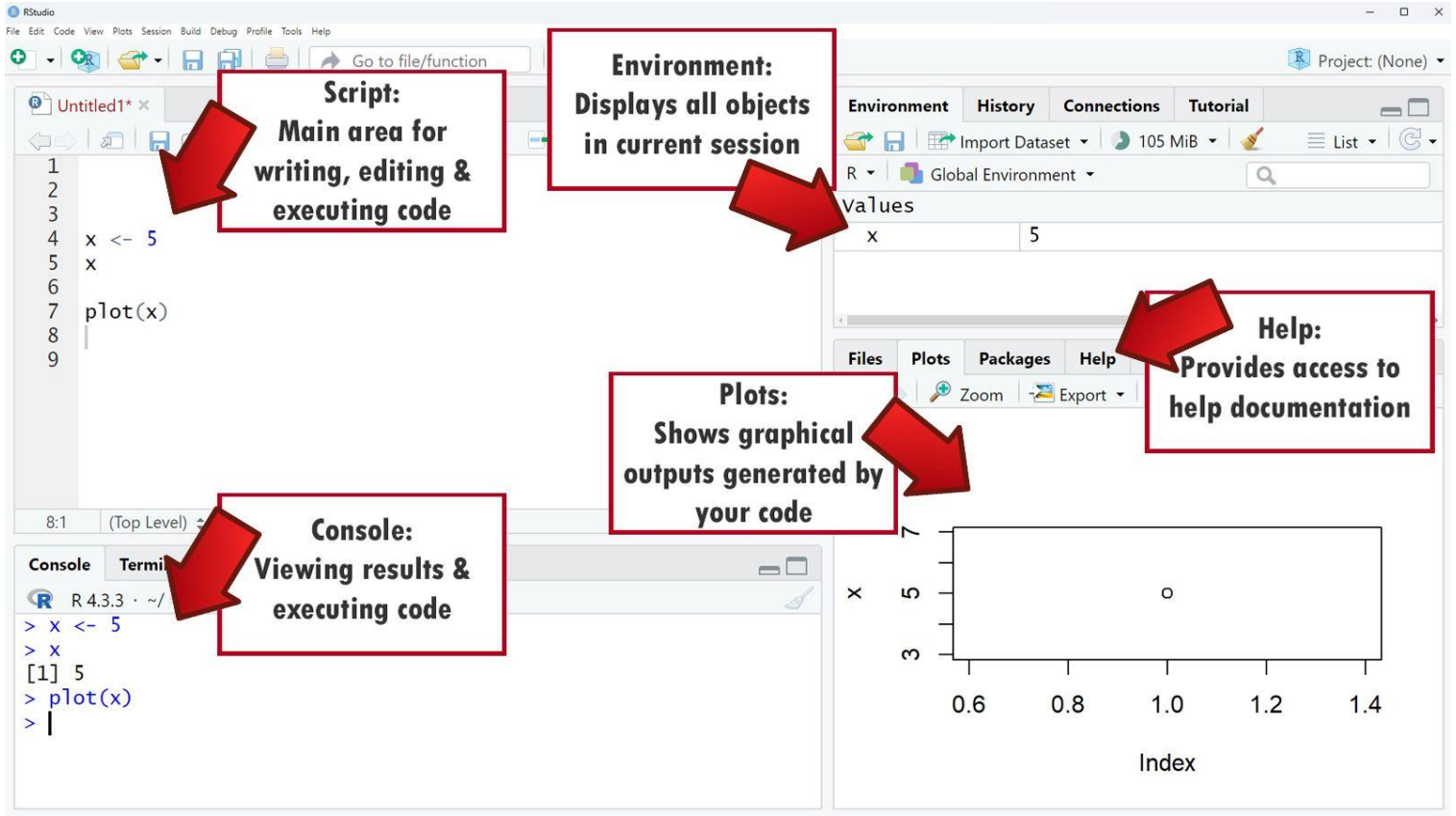
Boas práticas no R

- Área de trabalho do R Studio (painéis)



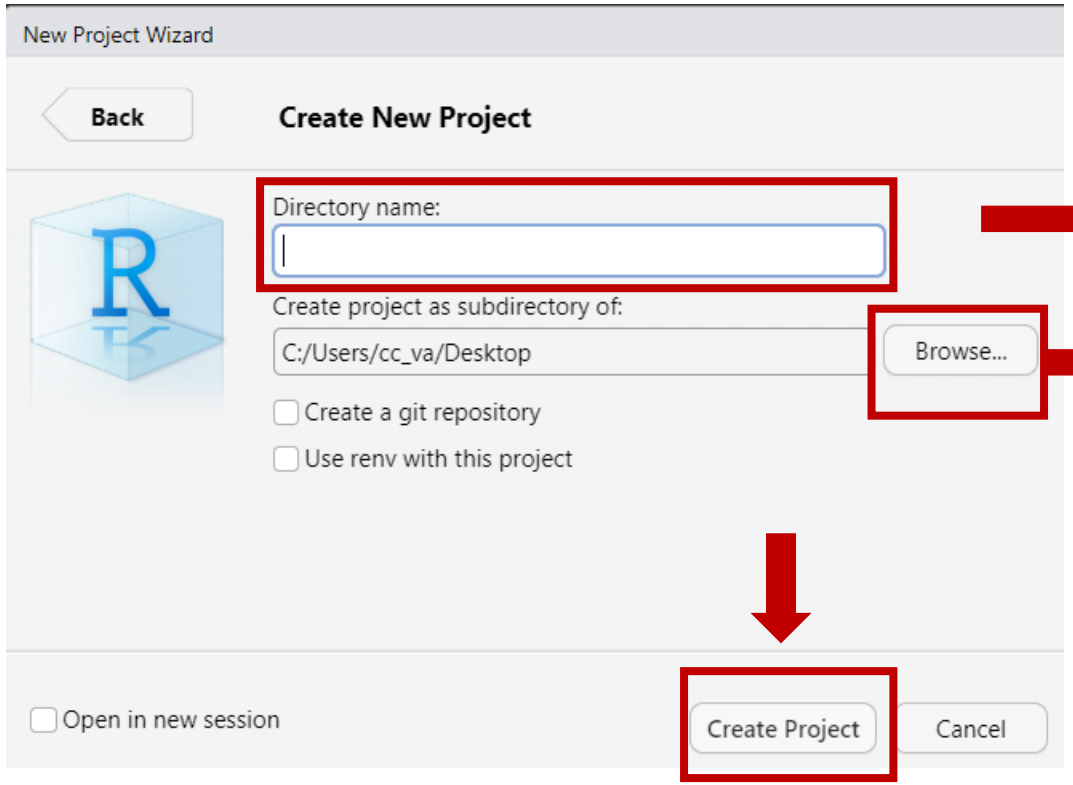
Boas práticas no R

- Área de trabalho do R Studio (o que você precisa saber!)



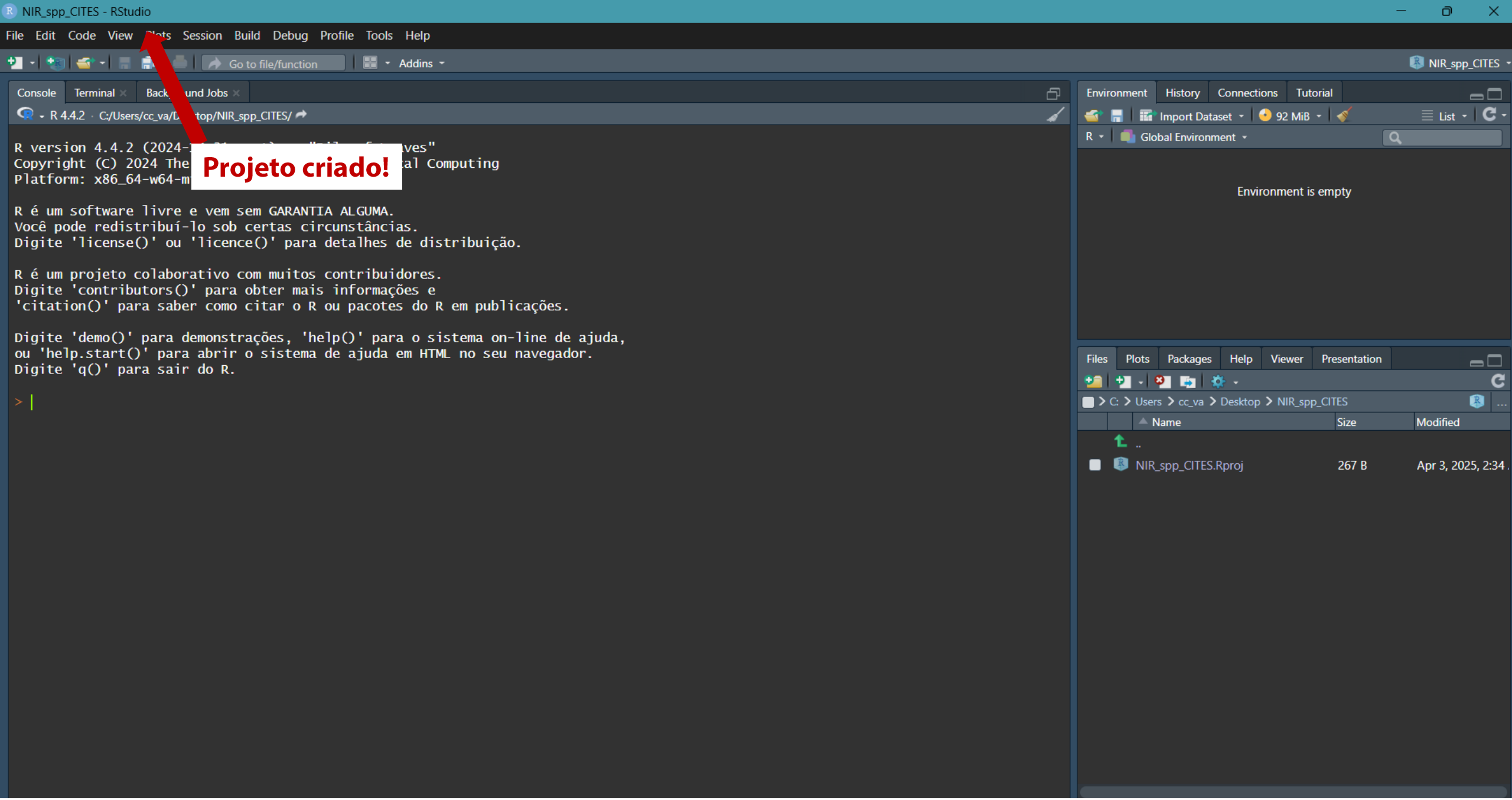
Boas práticas no R

- Criando um projeto no R Studio (.Rproj)
- File > New Project > New Directory > New Project



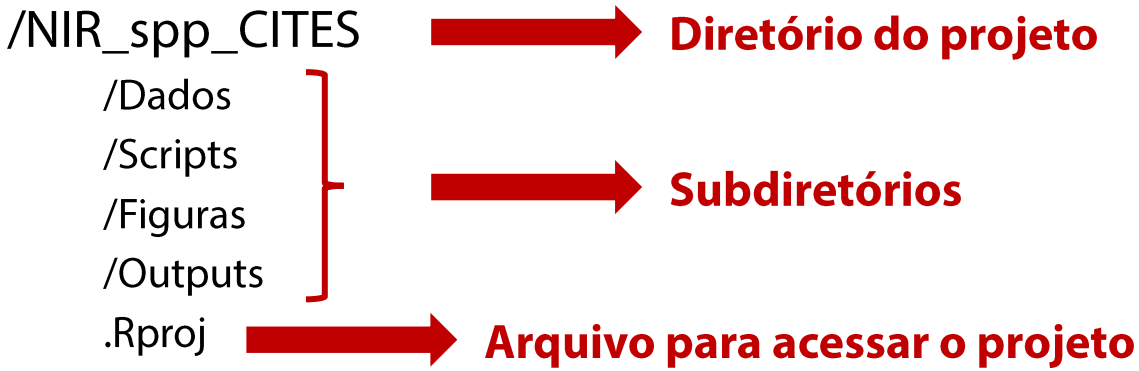
Vamos usar "NIR_spp_CITES"

Define o local onde a pasta deve ficar. No meu caso eu deixei no Desktop


NIR_spp_CITES - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
NIR_spp_CITES
Console
Terminal
Background Jobs
R 4.4.2 · C:/Users/cc_va/Desktop/NIR_spp_CITES/
Projeto criado!
R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.
R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.
Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.
Environment
History
Connections
Tutorial
Import Dataset
92 MiB
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
Presentation
C: > Users > cc_va > Desktop > NIR_spp_CITES
Name
Size
Modified
NIR_spp_CITES.Rproj
267 B
Apr 3, 2025, 2:34

Boas práticas no R

- Estruturas de pastas

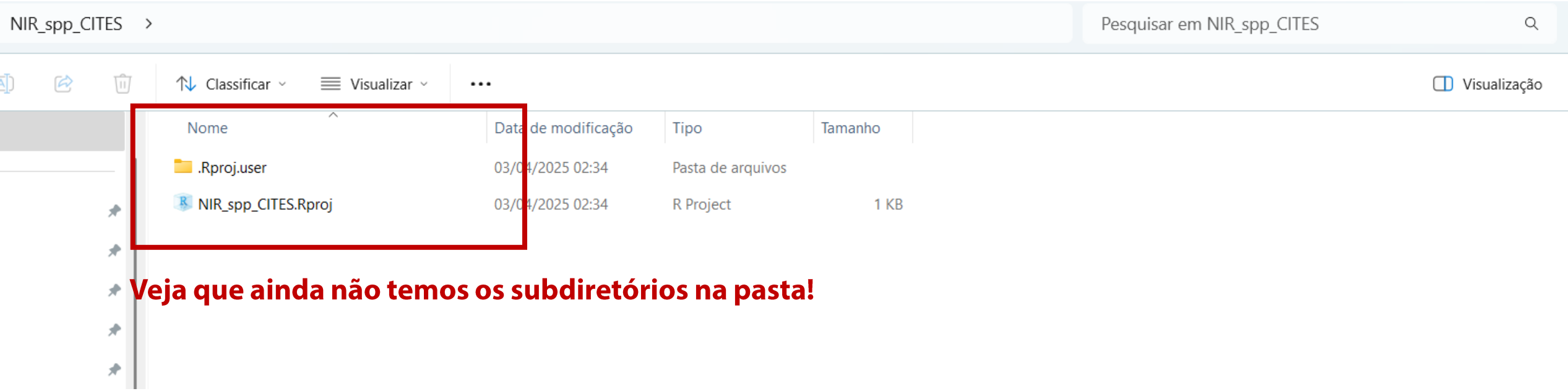


- Agora vamos, criar esses subdiretórios!

Boas práticas no R

- Estruturas de pastas

/NIR_spp_CITES
.Rproj



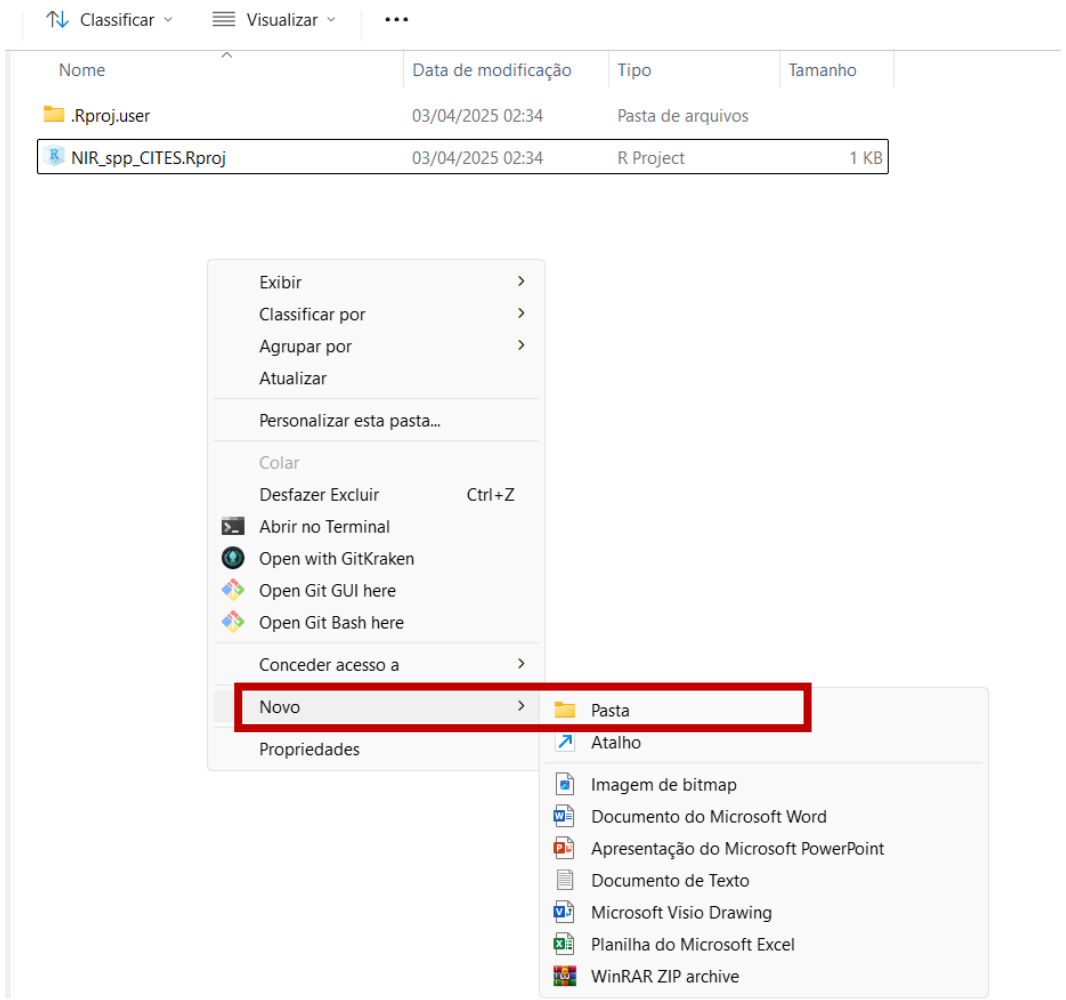
Veja que ainda não temos os subdiretórios na pasta!

Boas práticas no R

Você pode fazer de forma manual

Ou

Através do R Studio (próximos slides)



Faça isso para cada:

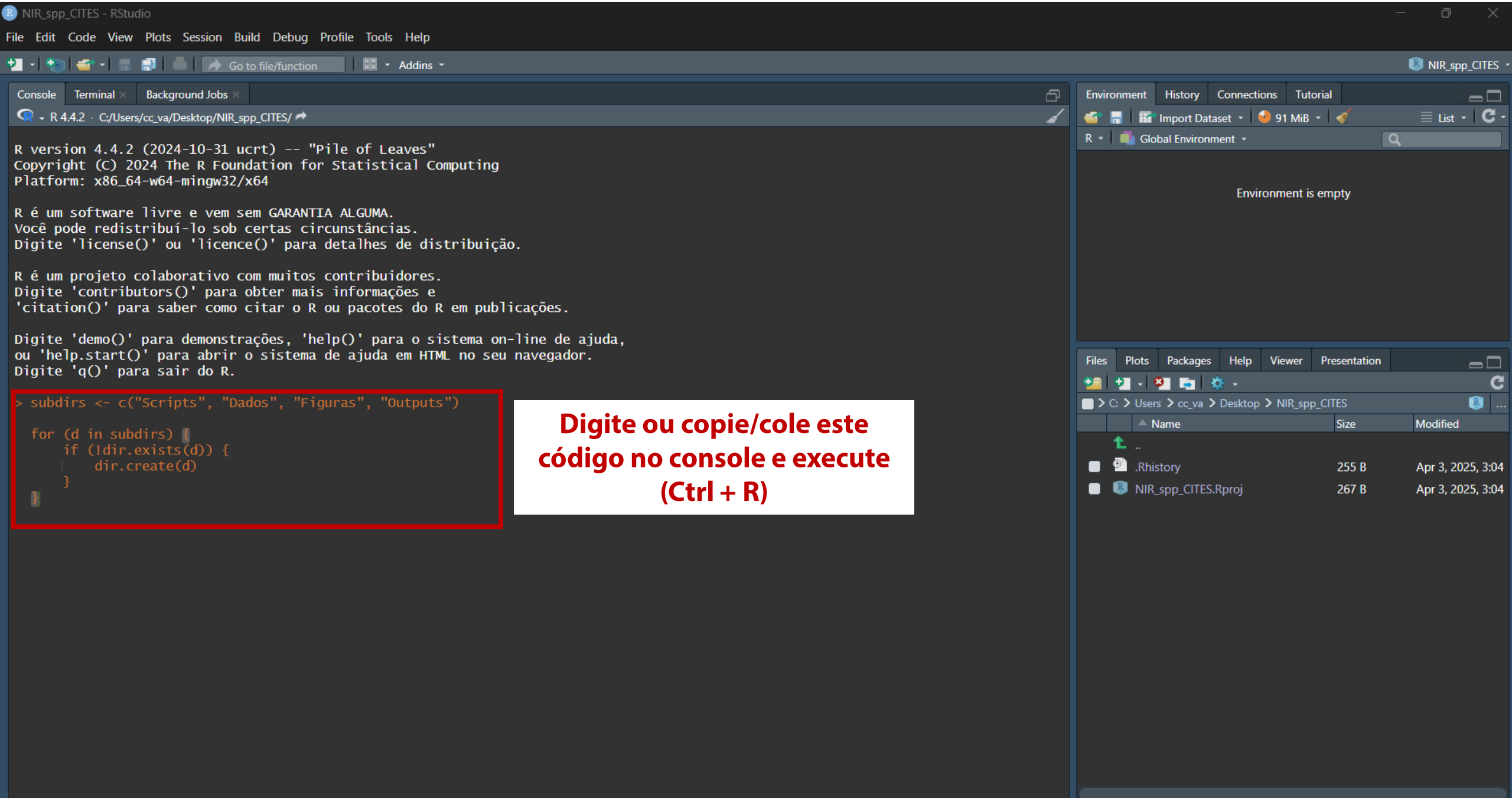
/Dados
/Scripts
/Figuras
/Outputs

Boas práticas no R

- Código para copiar e colar no **console** do R Studio:

```
subdirs <- c("Scripts", "Dados", "Figuras", "Outputs")

for (d in subdirs) {
  if (!dir.exists(d)) {
    dir.create(d)
  }
}
```


NIR_spp_CITES - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
NIR_spp_CITES
Console
Terminal
Background Jobs
R 4.4.2 · C:/Users/cc_va/Desktop/NIR_spp_CITES/
R version 4.4.2 (2024-10-31 ucrt) -- "Pile of Leaves"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64
R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.
R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.
Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.
> subdirs <- c("Scripts", "Dados", "Figuras", "Outputs")
for (d in subdirs) {
if (!dir.exists(d)) {
dir.create(d)
}
}
Digite ou copie/cole este código no console e execute (Ctrl + R)
Environment
History
Connections
Tutorial
Import Dataset
91 MiB
List
R
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
Presentation
C: > Users > cc_va > Desktop > NIR_spp_CITES
Name
Size
Modified
..
.Rhistory
255 B
Apr 3, 2025, 3:04
NIR_spp_CITES.Rproj
267 B
Apr 3, 2025, 3:04

Executando um script R

- Copie os arquivos das leituras espectrais para o subdiretório **“Dados”**
- Copie também o arquivo “**Handroanthus.xlsx**” para o subdiretório **“Dados”**
- Agora, copie o arquivo **“01_Importando e preparando dados espectrais.R”** para o subdiretório **“Scripts”** na pasta do projeto
- Com o projeto aberto, abra o script no R Studio: File > Open File > Selecione a pasta **Scripts** > Selecione o arquivo do script **“01_Importando e preparando dados espectrais.R”**

Acompanhe a execução do script na tela do R Studio

Agora é com vocês:

Mãos à obra!

Lista de presença (Dia 2/tarde)

https://forms.gle/WnXsJYpzR1aw5kxT8

ou