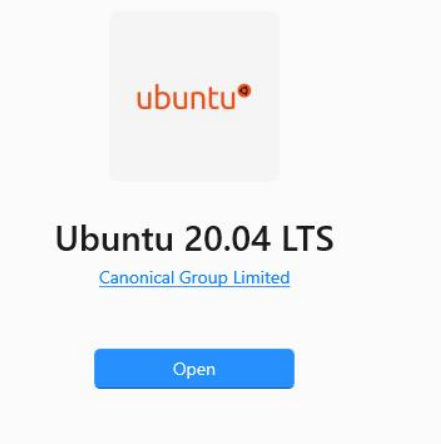
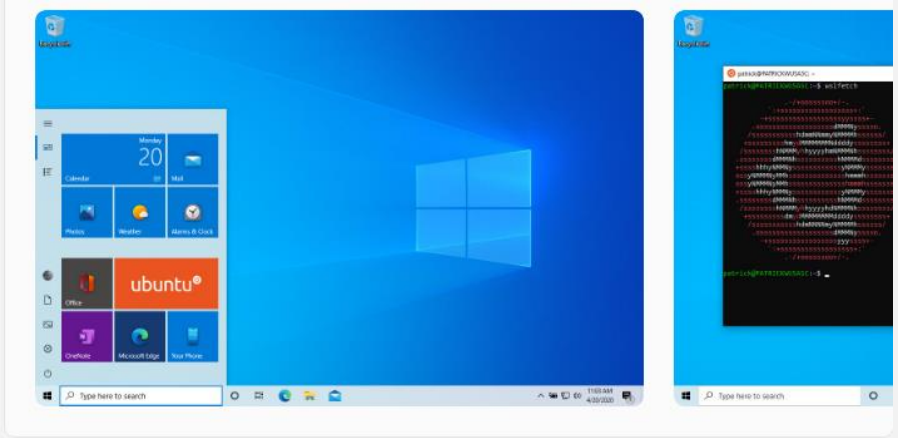
COMPUTER RESOURCES

- Computer

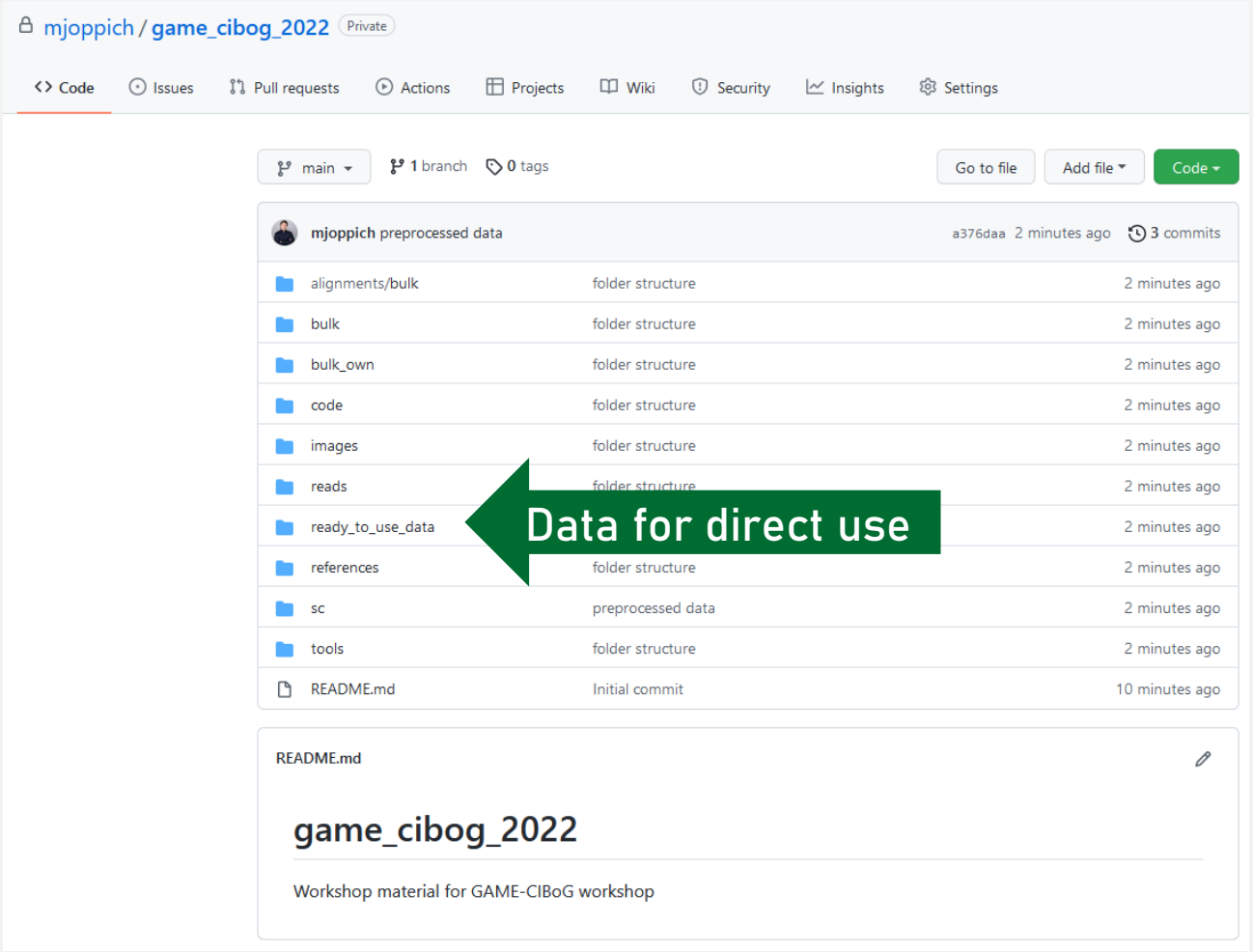
- Some kind of Linux
  - You use Linux? → you're fine (but need Win/Mac to run Loupe browser)
  - You use Windows? → use Windows Subsystem for Linux → you're fine





  - You use Mac OS? → you're mostly fine, but I can't help you
    - My Mac experience dates back to Mac OS X Leopard!

GIT REPOSITORY



- git clone https://github.com/mjoppich/game_cibog_2022.git
- Folder structure useful (not required though!) for workshop
- Send eMail to

  joppich@bio.ifi.lmu.de

  with Github username to get access

SOFTWARE RESOURCES

- You need some R version

```
mjoppich@spectre3:/mnt/c/Users/mjopp$ R --version
R version 4.0.3 (2020-10-10) -- "Bunny-Wunnies Freak Out"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)
```

- You may want to use jupyter+R kernel
  - `install.packages("devtools")`
  - `devtools::install_github("IRkernel/IRkernel")`
  - `IRkernel::installspec()`

- Some kind of IDE or jupyter-lab in the browser
  - Big fan of Visual Code
    - View PDF, PNG, JPEG
    - Supports jupyter notebooks and plots

```
options(jupyter.plot_mimetypes = 'image/png')
options(repr.plot.width = 1, repr.plot.height = 0.75, repr.plot.res = 100)
```

REQUIRED TOOLS FOR DOWNLOADING DATA

Getting the read data

- sratoolkit for downloading files
  - wget --output-document tools/sratoolkit.tar.gz http://ftp-trace.ncbi.nlm.nih.gov/sra/sdk/current/sratoolkit.current-ubuntu64.tar.gz
  - cd tools && tar -vxzf sratoolkit.tar.gz
  - vdb-config –interactive

  - Make path to sratoolkit visible:
  - Add

    export PATH=$PATH:/mnt/t/workshop/tools/sratoolkit.2.11.2-ubuntu64/bin

    to ~/.bashrc and `source ~/.bashrc`

- Then we can call *fastq-dump* as

  `fastq-dump --gzip SRR9048093`

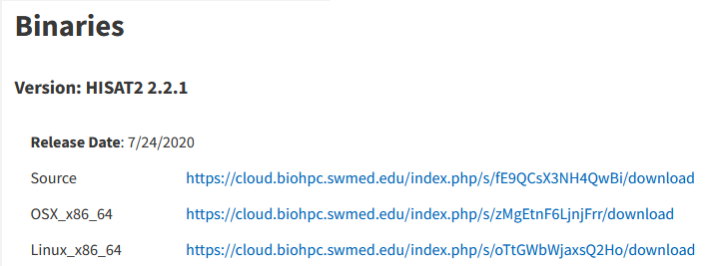
  for downloading specific SRA-files

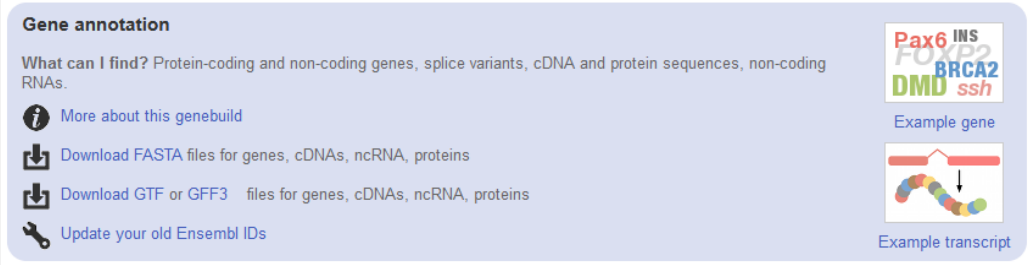
REQUIRED FILES AND REFERENCES

Doing the alignment

- Hisat2
  - `sudo apt install hisat2`
  - http://daehwankimlab.github.io/hisat2/
  - Index/Reference genome: H. sapiens, GRCh38, https://genome-idx.s3.amazonaws.com/hisat/grch38_genome.tar.gz
  - `tar xfz grch38_genome.tar.gz`
- featureCounts / subread
  - `sudo apt install subread`
  - https://sourceforge.net/projects/subread/files/subread-2.0.3/

  - Genome Annotation File:
    - http://ftp.ensembl.org/pub/release-105/gtf/homo_sapiens/Homo_sapiens.GRCh38.105.gtf.gz
    - `gunzip Homo_sapiens.GRCh38.105.gtf.gz`

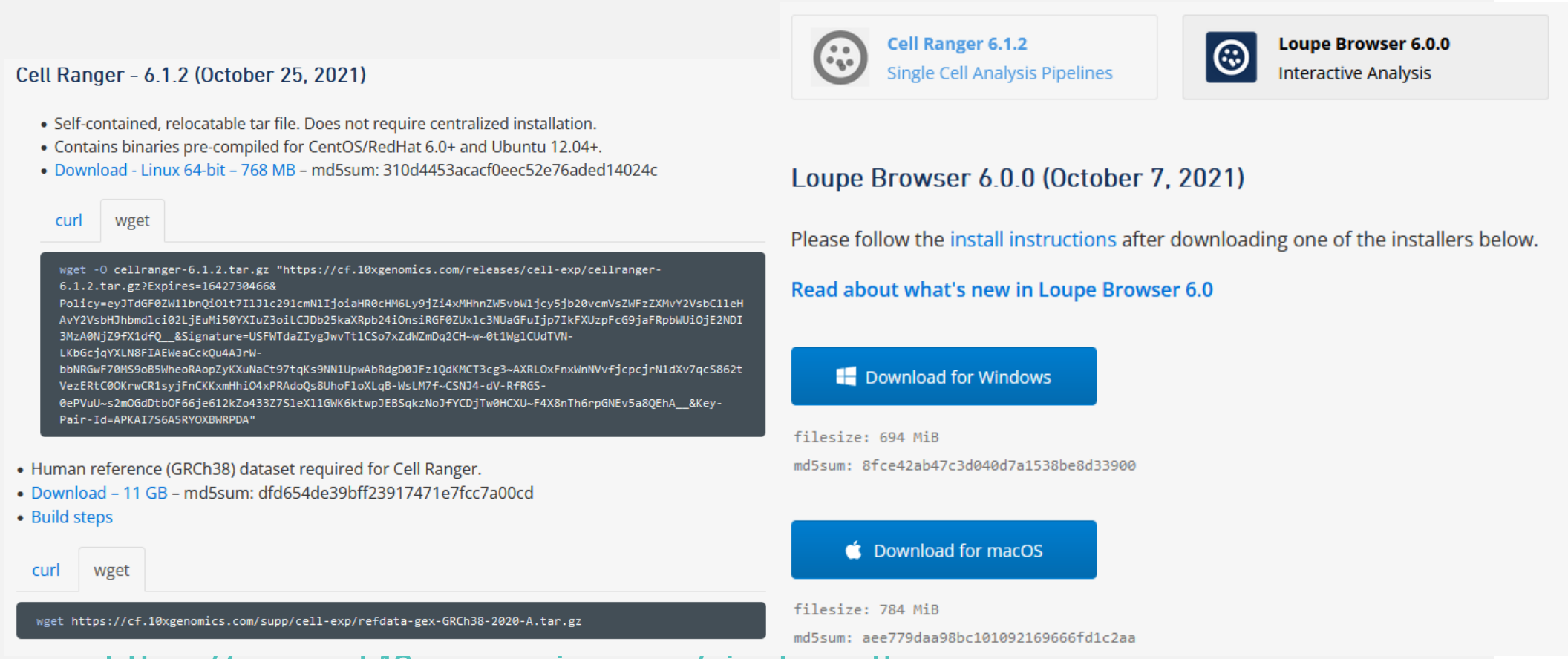




R PACKAGES

```
# ensures we can install everything
library(devtools) # => install.packages("devtools")


# accessing files, plotting
library(readr) # => install.packages("readr")
library(ggplot2) # => install.packages("ggplot2")
library(cowplot) # => install.packages("cowplot")
library(ggrepel) # => install.packages("ggrepel")
library(pheatmap) # => install.packages("pheatmap")
library(RColorBrewer) # => install.packages(" RColorBrewer ")


# Bioconductor package manager => required for basically all bioinformatics tasks
if (!requireNamespace('BiocManager', quietly = TRUE))
  install.packages('BiocManager')

library(EnhancedVolcano) # => BiocManager::install('EnhancedVolcano') or devtools::install_github('kevinblighe/EnhancedVolcano'
library(DESeq2) # => BiocManager::install('DESeq2')
library(clusterProfiler) # => BiocManager::install('clusterProfiler')
library(DOSE) # => BiocManager::install('DOSE')
library(enrichplot) # => BiocManager::install('enrichplot')
library(ReactomePA) # => BiocManager::install('ReactomePA')
library(org.Hs.eg.db) # => BiocManager::install('org.Hs.eg.db')

library(disgenet2r) # => devtools::install_bitbucket("ibi_group/disgenet2r")
```

DOWNLOADING CELLRANGER+LOUPE



Cell Ranger - 6.1.2 (October 25, 2021)

- Self-contained, relocatable tar file. Does not require centralized installation.
- Contains binaries pre-compiled for CentOS/RedHat 6.0+ and Ubuntu 12.04+.
- Download - Linux 64-bit – 768 MB – md5sum: 310d4453acacf0eec52e76aded14024c

curl wget

```
wget -O cellranger-6.1.2.tar.gz "https://cf.10xgenomics.com/releases/cell-exp/cellranger-
6.1.2.tar.gz?Expires=1642730466&
Policy=eyJTdGF0ZW1lbnQiOlt7Il1lc291cmNlIjoiaHR0cHM6Ly9jZi4xMHhnZW5vbW1jcy5jb20vcmVsZWFzZXMvY2VsbC1leH
AvY2VsbHJhbmd1ci02LjEuMi50YXIuZ3oiLCJDb25kaXRpb24iOnsiRGF0ZUx1c3NUaGFuIjp7IkFXUzpFcG9jaFRpbWUiOjE2NDI
3MzA0NjZ9fX1dfQ__&Signature=USFWTdaZIygJwvTt1CSo7xZdWZmDq2CH~w~0t1WglCUdTVN-
LKbGcjqYXLN8FIAEWeaCckQu4AJrW-
bbNRGwF70MS9oB5WheoRAopZyKXuNaCt97tqKs9NN1UpwAbRdgD0JFz1QdKMCT3cg3~AXRLOxFnxWnNVvfjcpcjrN1dXv7qcS862t
VezERtC0OKrwCR1syjFnCKKxmHhiO4xPRAdoQs8UhoFloXLqB-WsLM7f~CSNJ4-dV-RfRGS-
0ePVuU~s2mOGdDtbOF66je612kZo433Z7SleX11GWK6ktwpJEBSqkzNoJfYCDjTw0HCXU~F4X8nTh6rpGNEv5a8QEhA__&Key-
Pair-Id=APKAI7S6A5RYOXBWRPDA"
```

- Human reference (GRCh38) dataset required for Cell Ranger.
- Download – 11 GB – md5sum: dfd654de39bff23917471e7fcc7a00cd
- Build steps

curl wget

```
wget https://cf.10xgenomics.com/supp/cell-exp/refdata-gex-GRCh38-2020-A.tar.gz
```

Cell Ranger 6.1.2
Single Cell Analysis Pipelines

Loupe Browser 6.0.0
Interactive Analysis

Loupe Browser 6.0.0 (October 7, 2021)

Please follow the install instructions after downloading one of the installers below.

Read about what's new in Loupe Browser 6.0

Download for Windows

filesize: 694 MiB

md5sum: 8fce42ab47c3d040d7a1538be8d33900

Download for macOS

filesize: 784 MiB

md5sum: aee779daa98bc101092169666fd1c2aa

- https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
- Might require brief registration

DOWNLOAD RSTUDIO

Deliver open source data science at *scale*

Learn how RStudio's enterprise-ready pro products help data science teams securely deploy their work, collaborate, and communicate with decision makers.

SEE OUR PRO PRODUCTS IN ACTION

RStudio Desktop 2021.09.2+382 - Release Notes

1. Install R. RStudio requires R 3.0.1+.

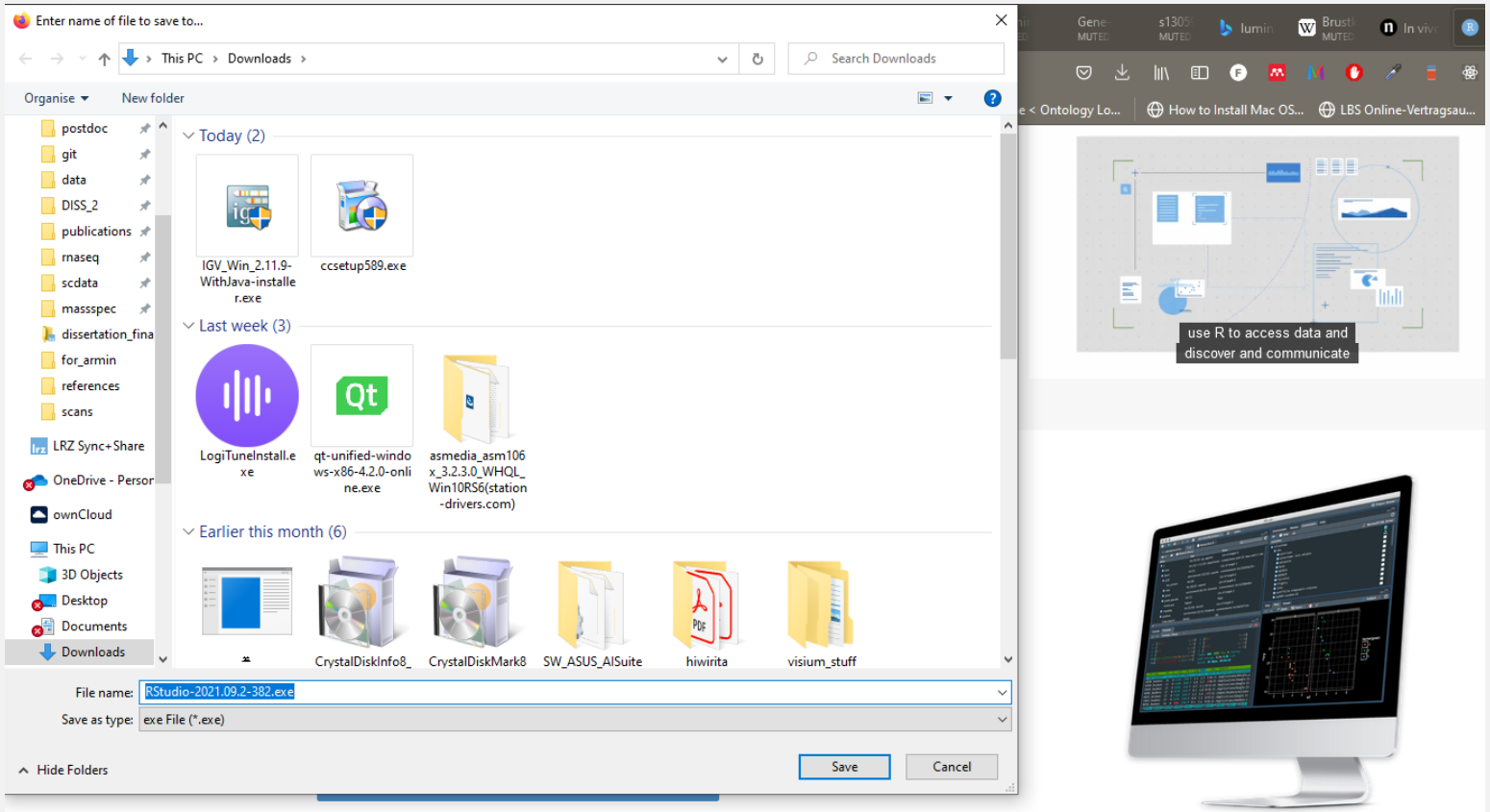
2. Download RStudio Desktop. Recommended for your system:

DOWNLOAD RSTUDIO FOR WINDOWS
2021.09.2+382 | 156.89MB

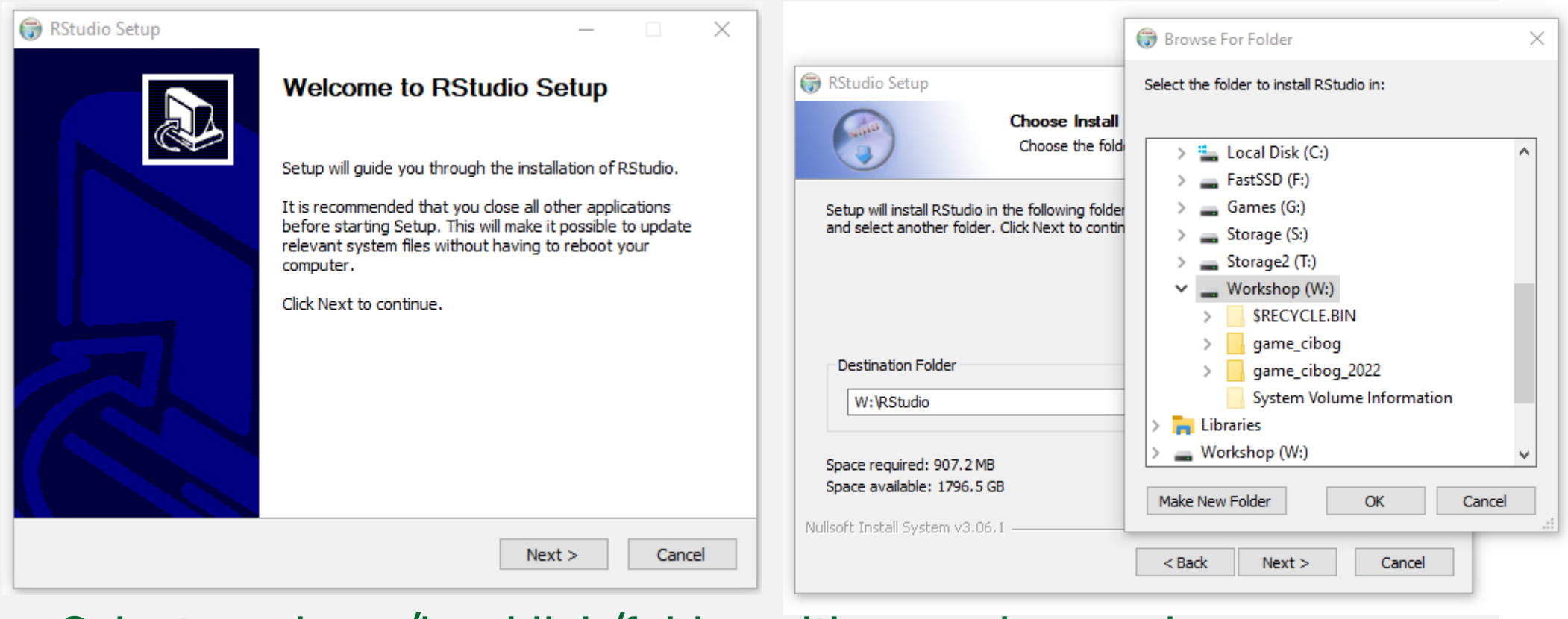
Requires Windows 10 (64-bit)

https://www.rstudio.com/products/rstudio/download/#download

DOWNLOAD RSTUDIO

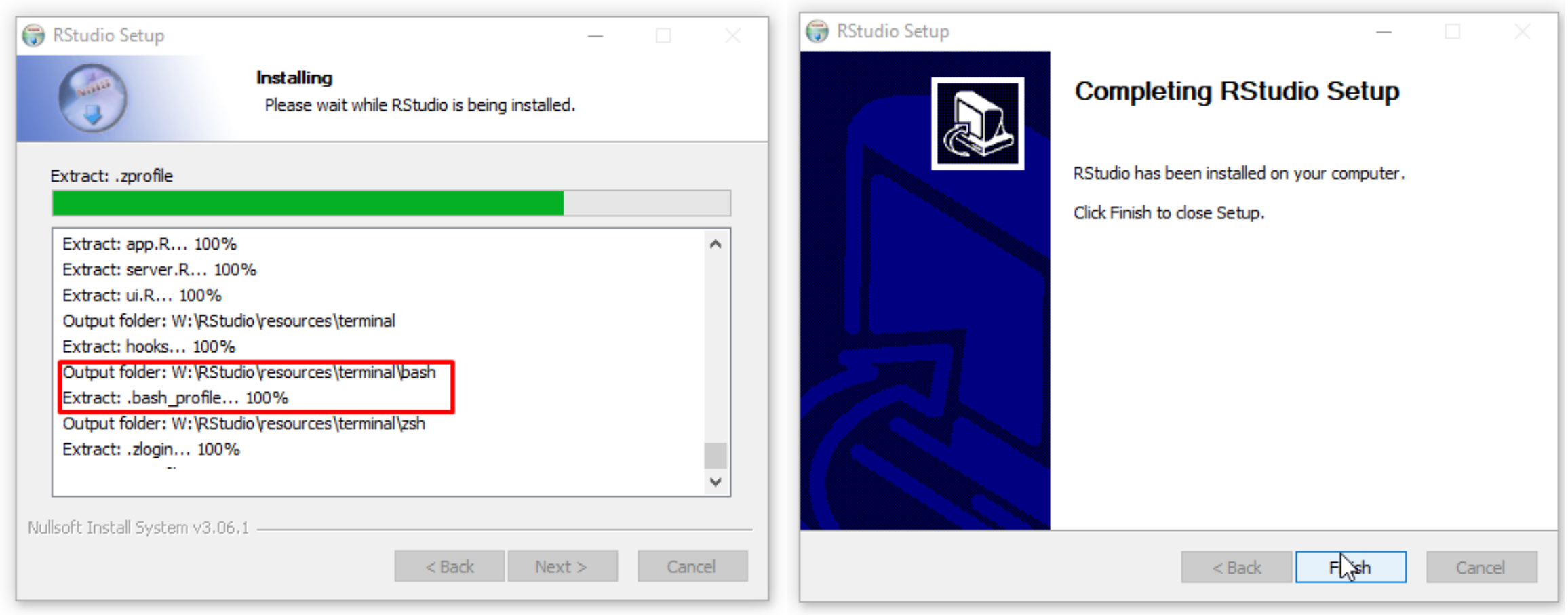


DOWNLOAD RSTUDIO

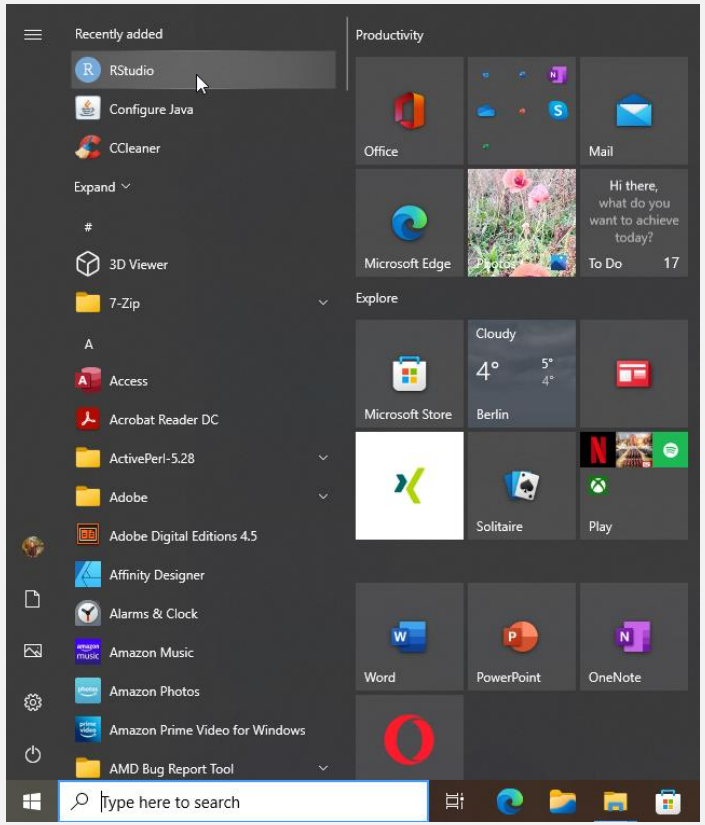


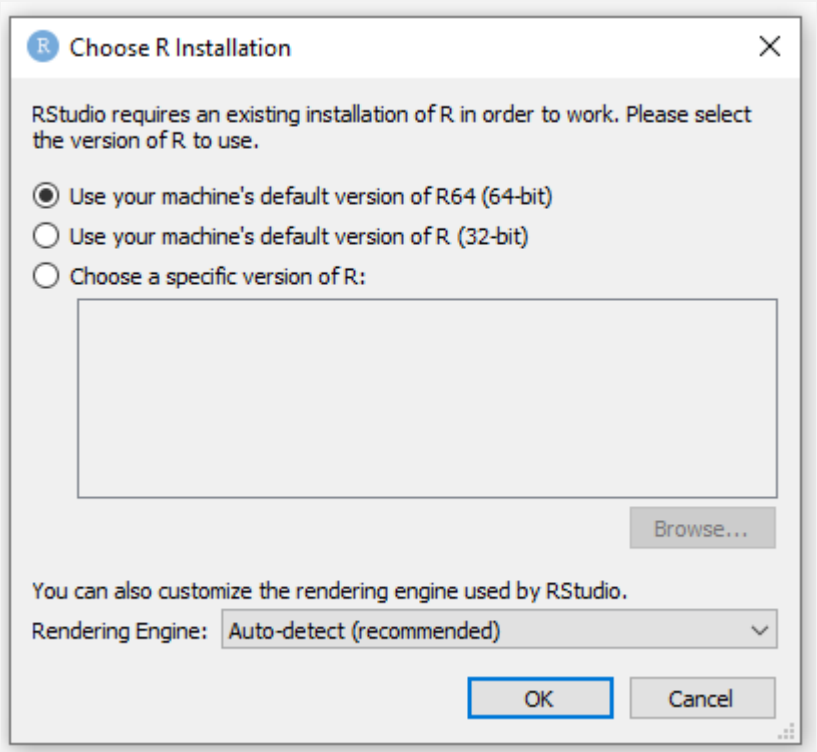
- Select a volume/harddisk/folder with enough space!

DOWNLOAD RSTUDIO



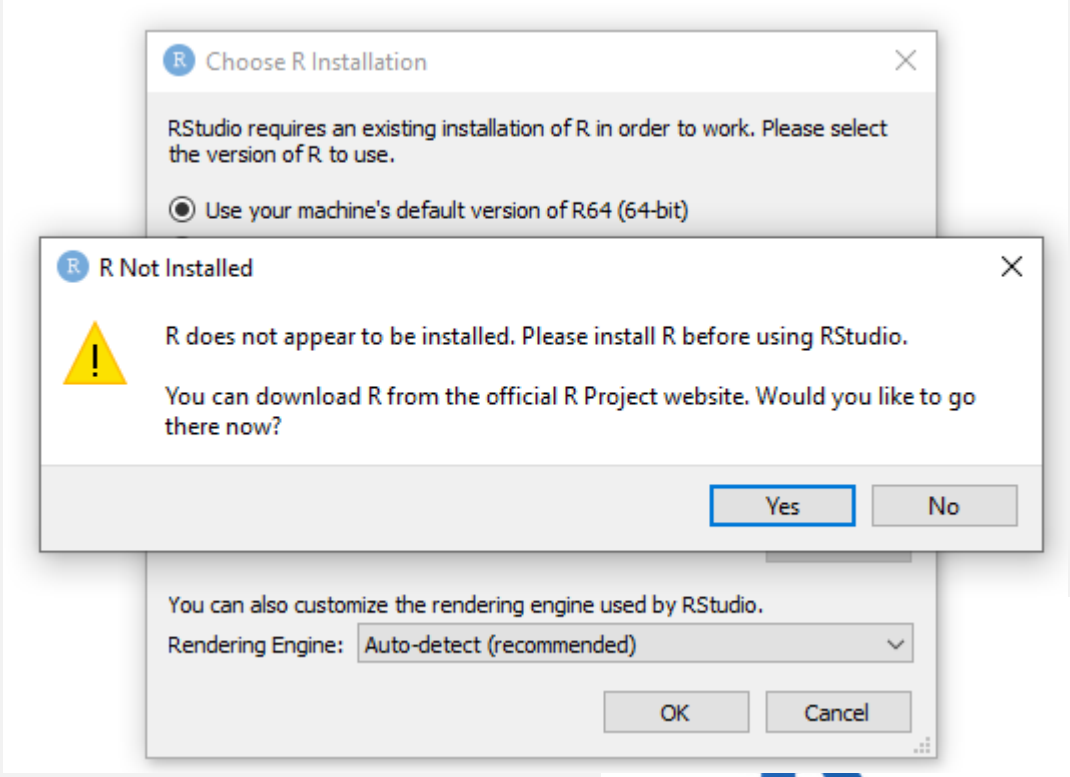
STARTING RSTUDIO

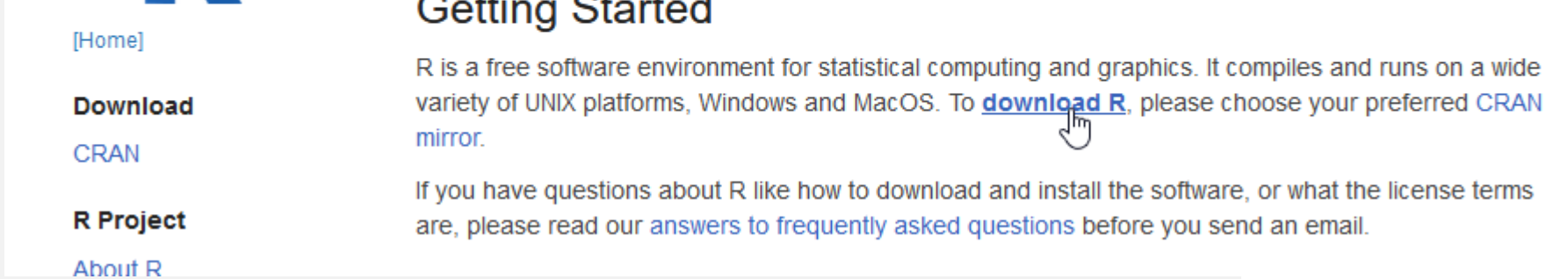




- When you start Rstudio the first time, select R version
  - If you have no R installed yet, select “Use your machine’s default version of R64”

INSTALLING R





DOWNLOADING R

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
Algeria	
https://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
https://mirror.aarnet.edu.au/pub/CRAN/	AARNET

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources

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 (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

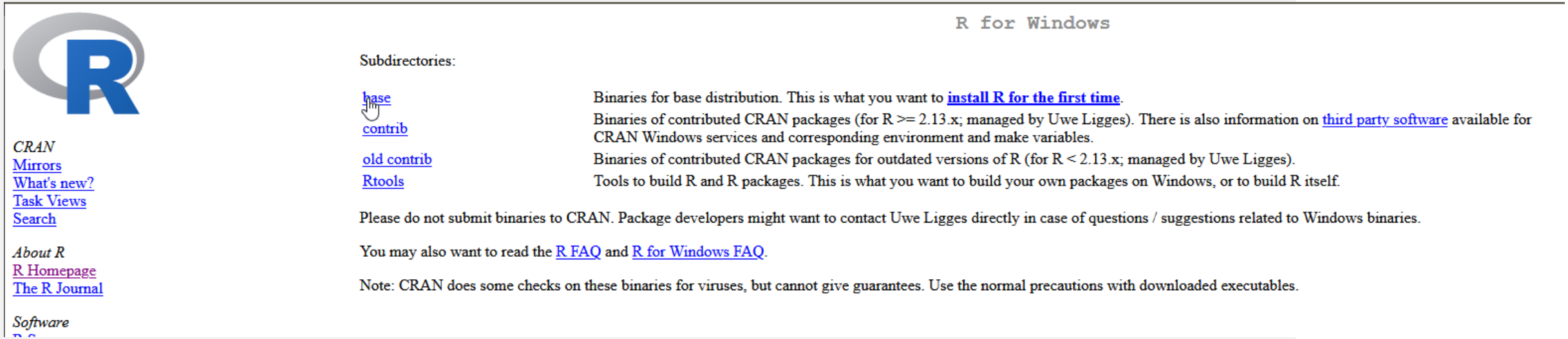
R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

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 (2021-11-01, Bird Hippie) R-4.1.2.tar.gz, read what's new in the latest version.

DOWNLOADING R



CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software

R for Windows

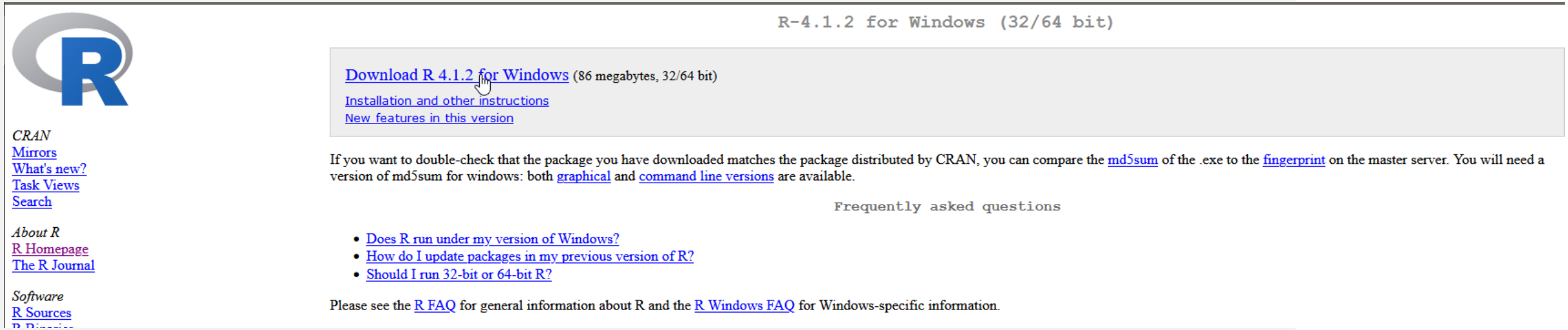
Subdirectories:

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).
Rtools	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.



CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources

R-4.1.2 for Windows (32/64 bit)

Download R 4.1.2 for Windows (86 megabytes, 32/64 bit)
Installation and other instructions
New features in this version

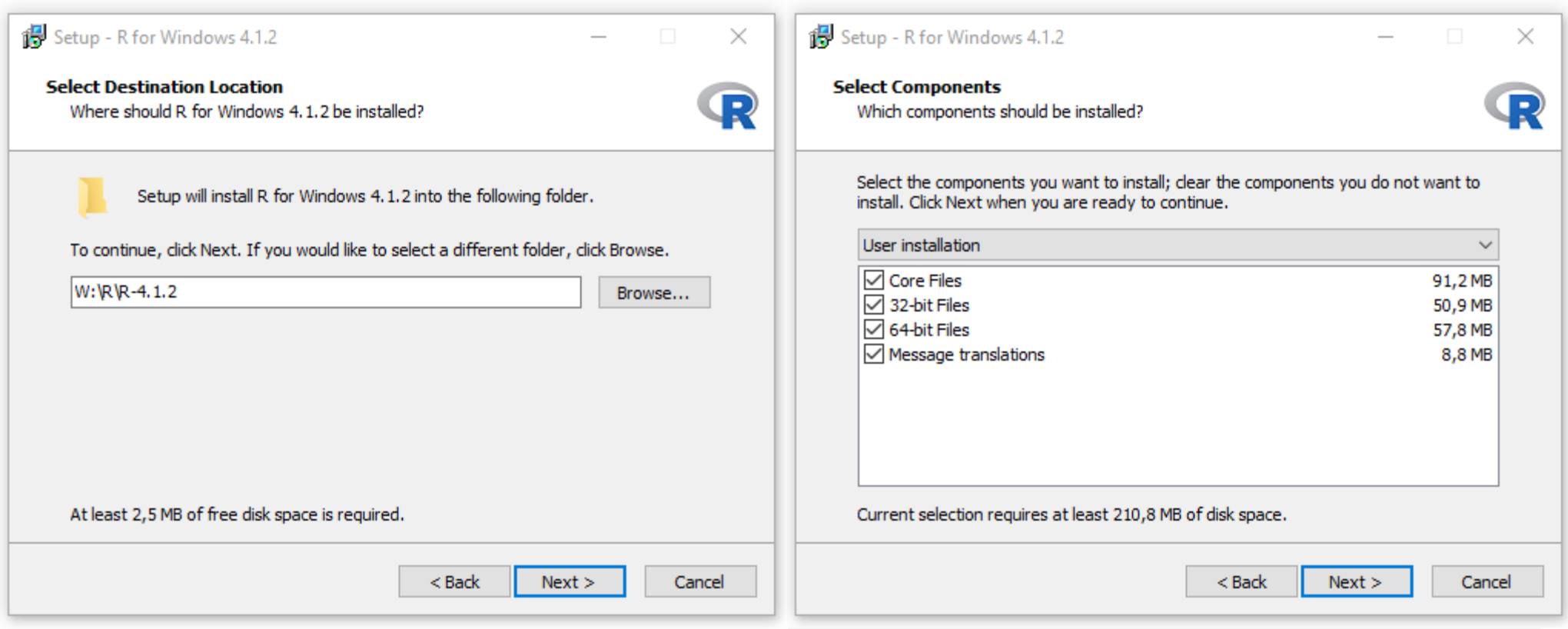
If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the md5sum of the .exe to the fingerprint on the master server. You will need a version of md5sum for windows: both graphical and command line versions are available.

Frequently asked questions

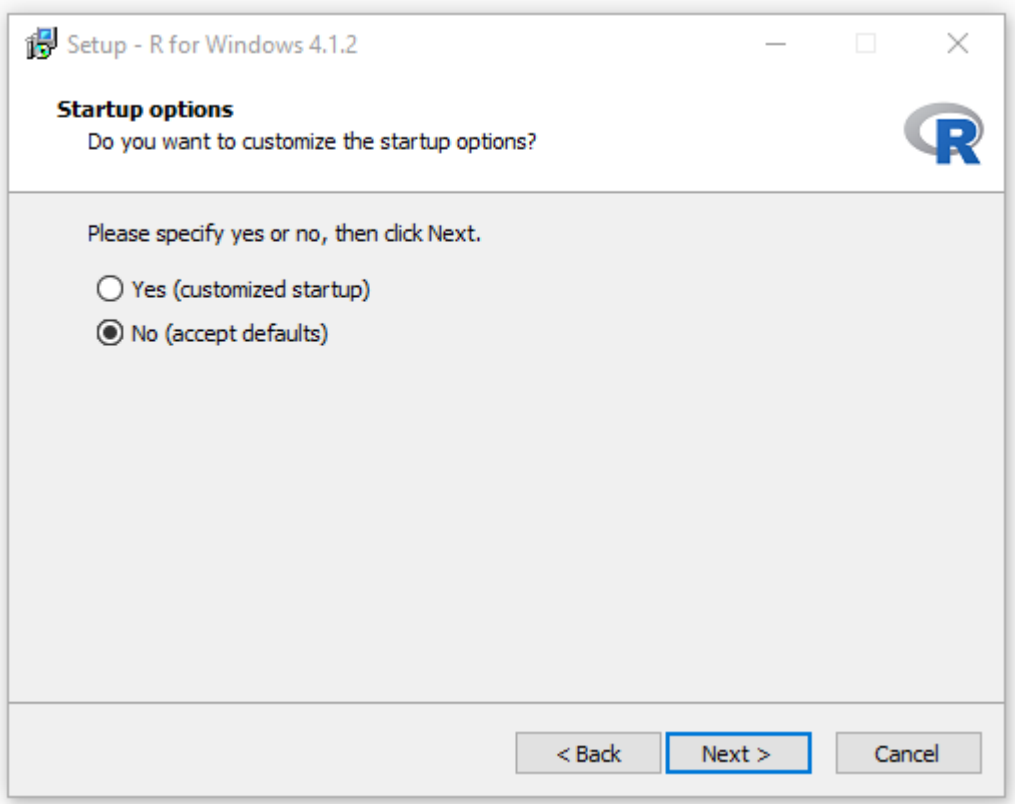
- Does R run under my version of Windows?
- How do I update packages in my previous version of R?
- Should I run 32-bit or 64-bit R?

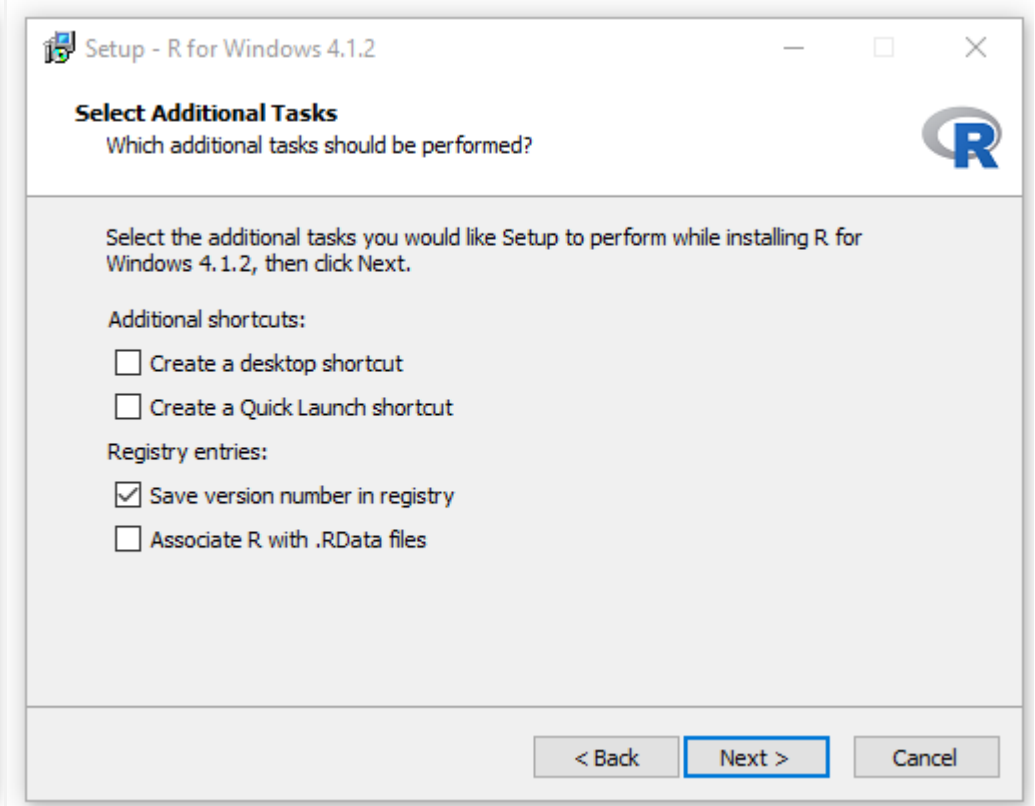
Please see the R FAQ for general information about R and the R Windows FAQ for Windows-specific information.

INSTALLING R

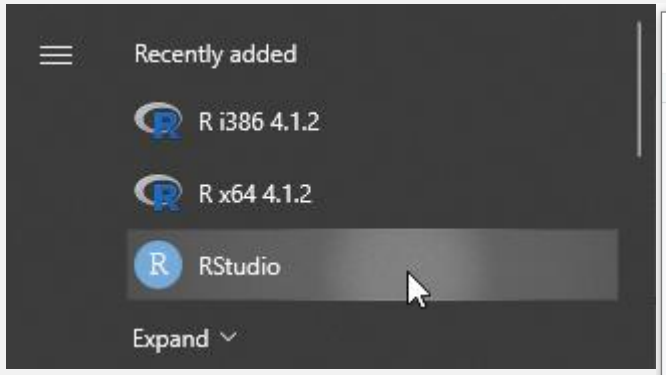


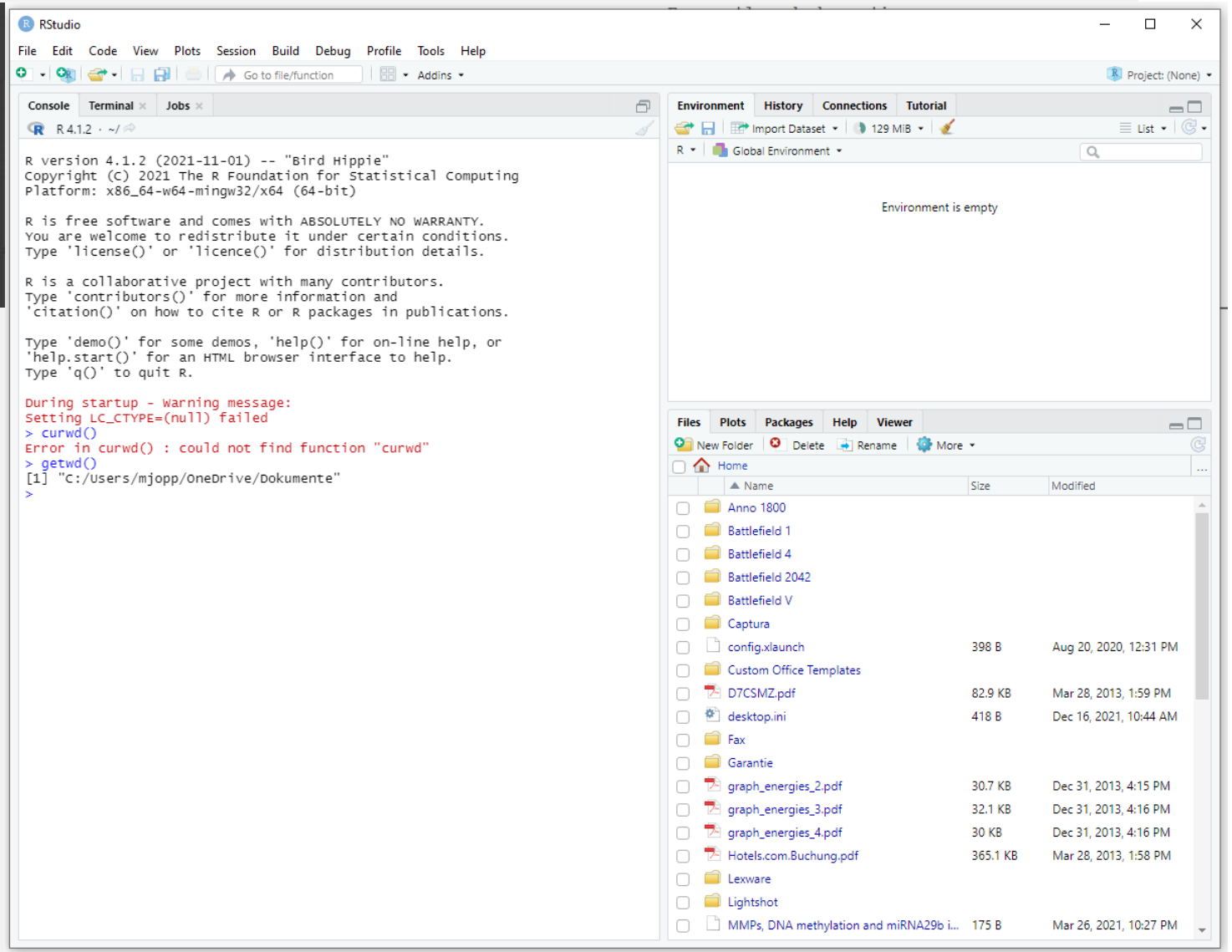
INSTALLING R





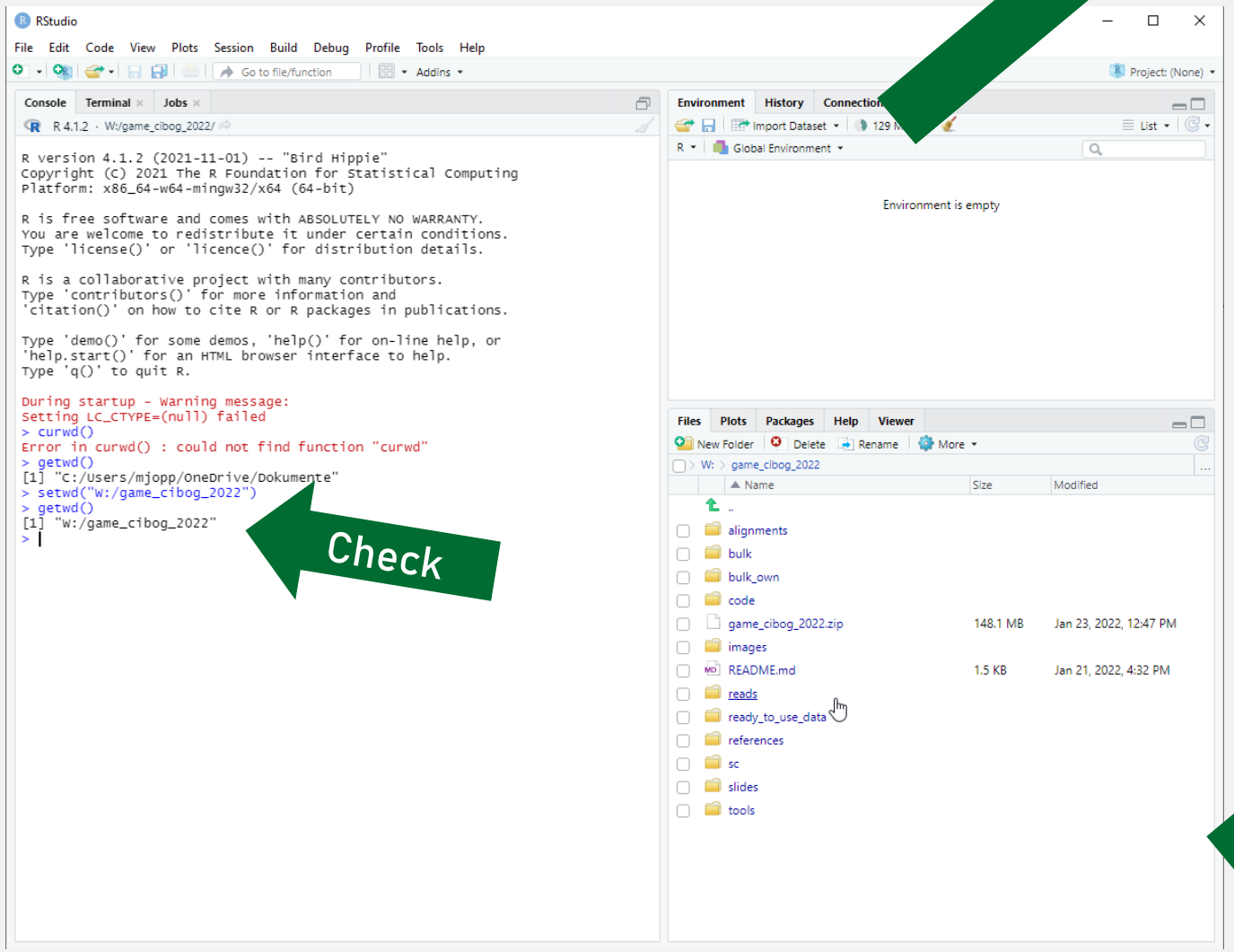
STARTING RSTUDIO

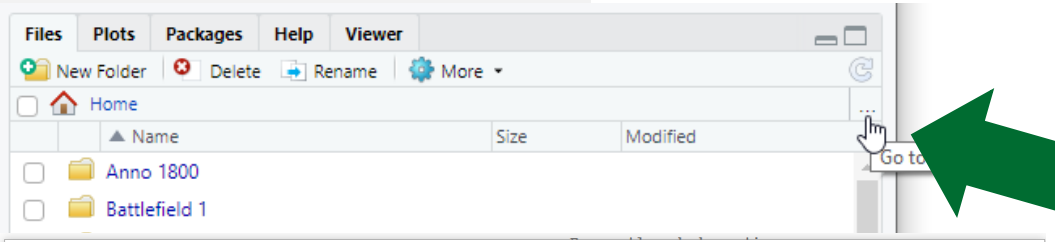


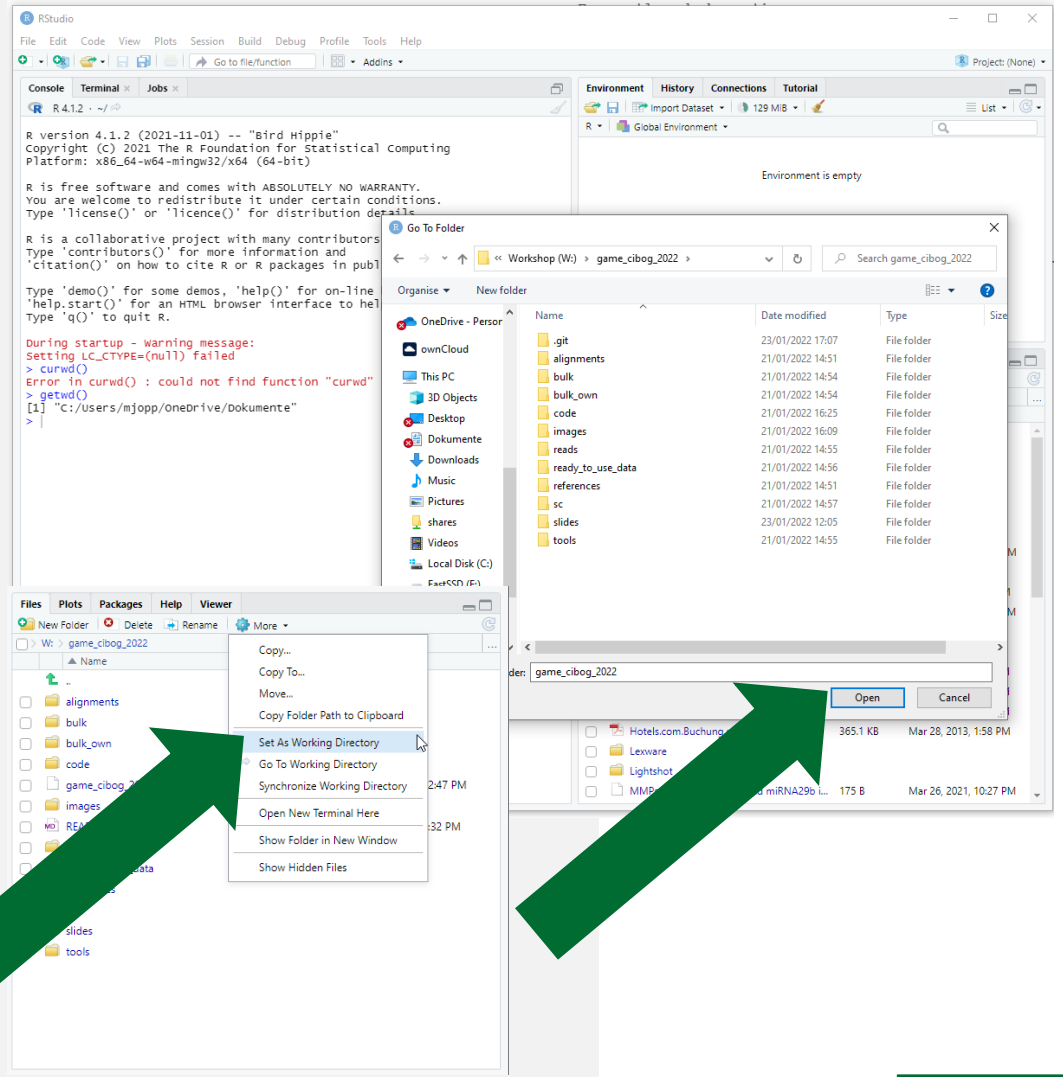


STARTING RSTUDIO

Changing the current working directory

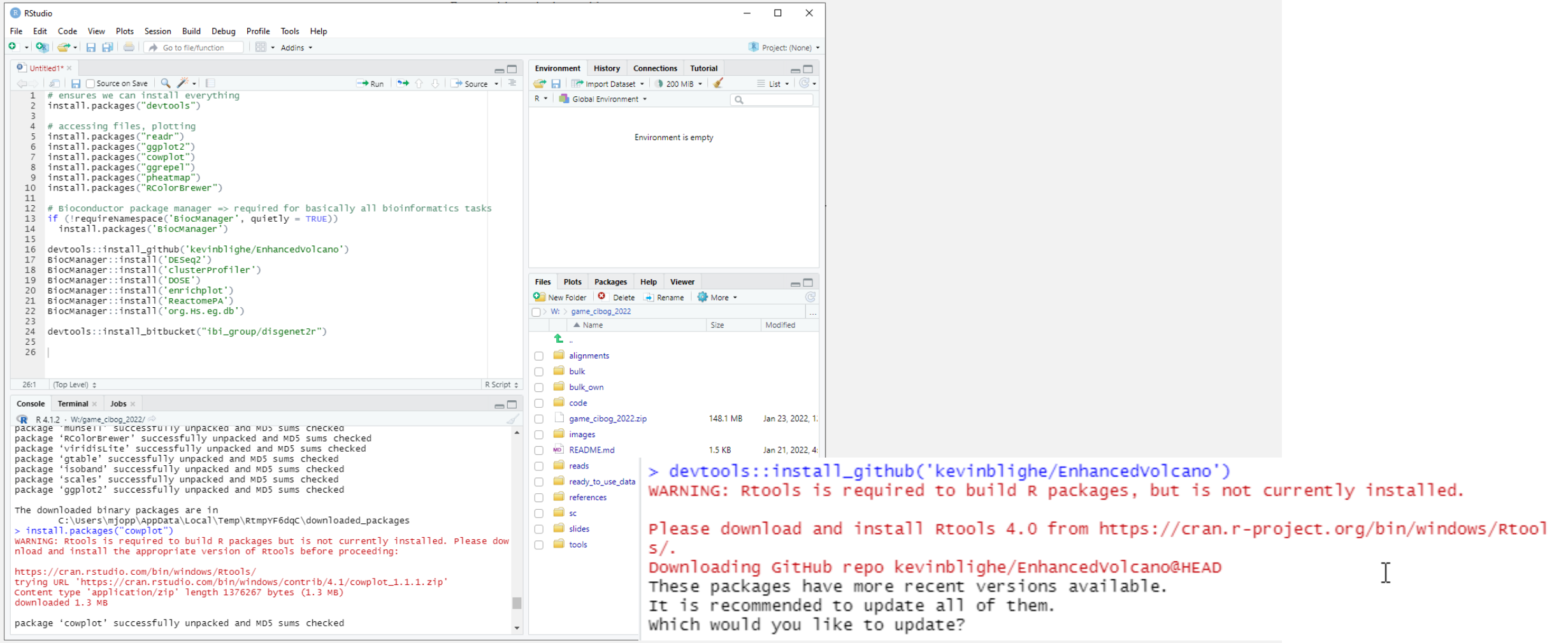




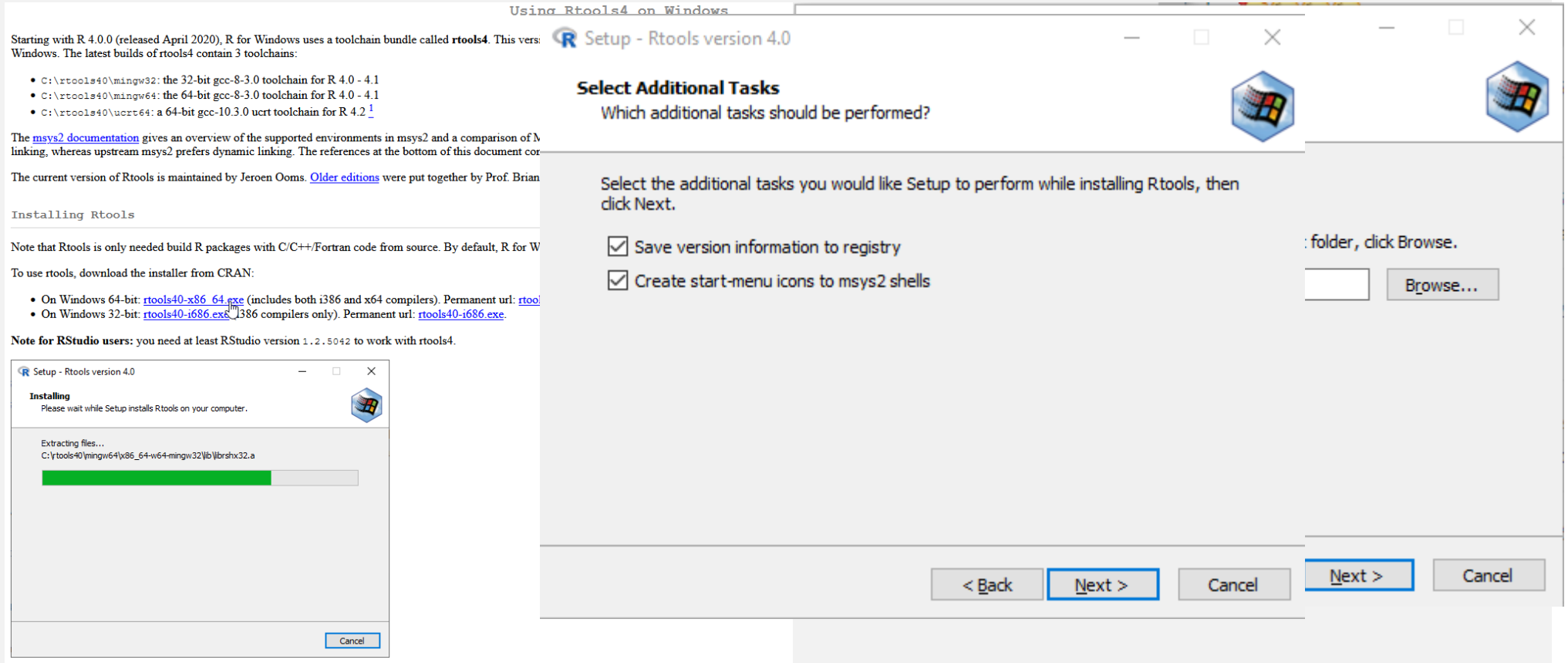


RTOOLS REQUIRED

Some libraries require compilation from source. On Windows: Rtools required

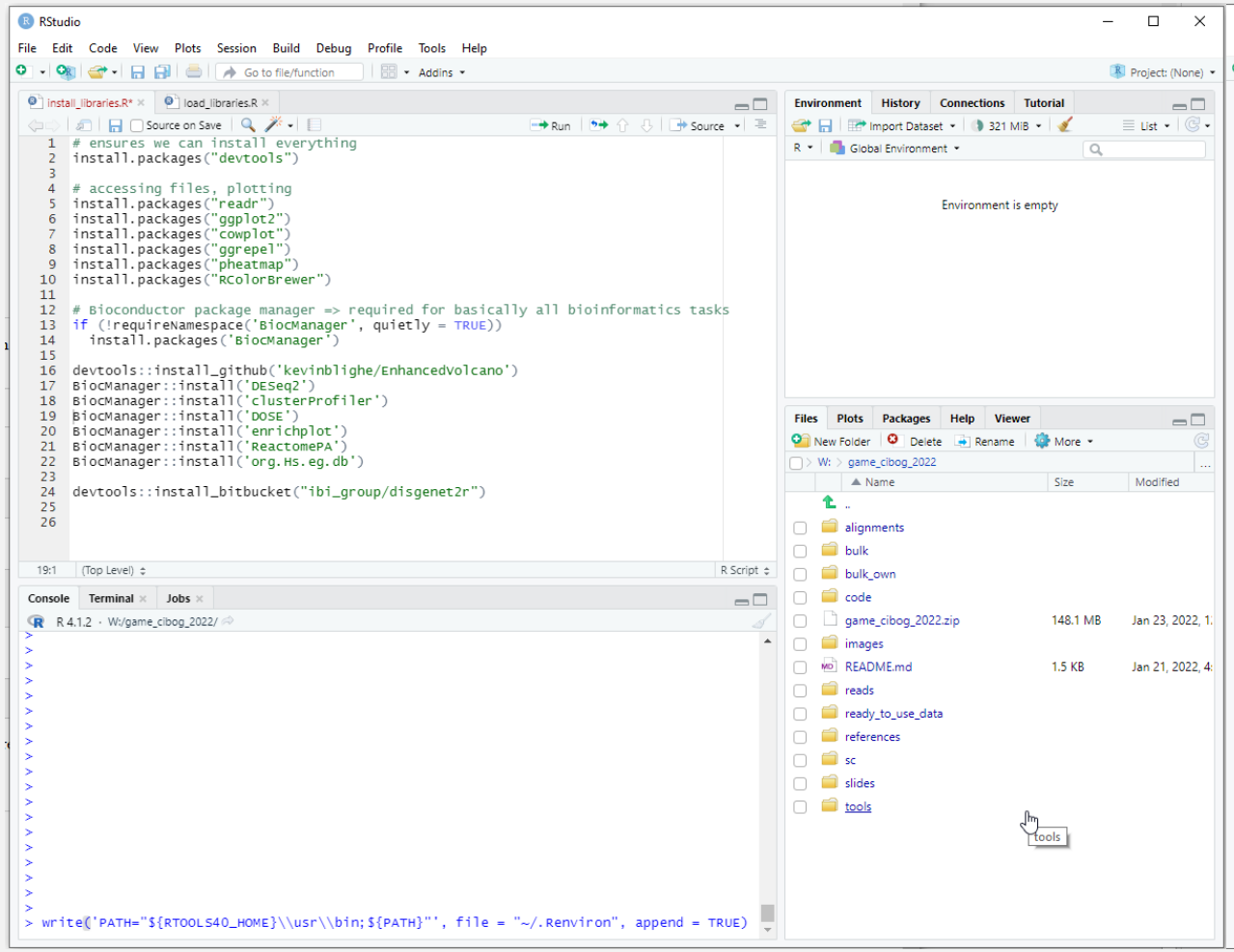


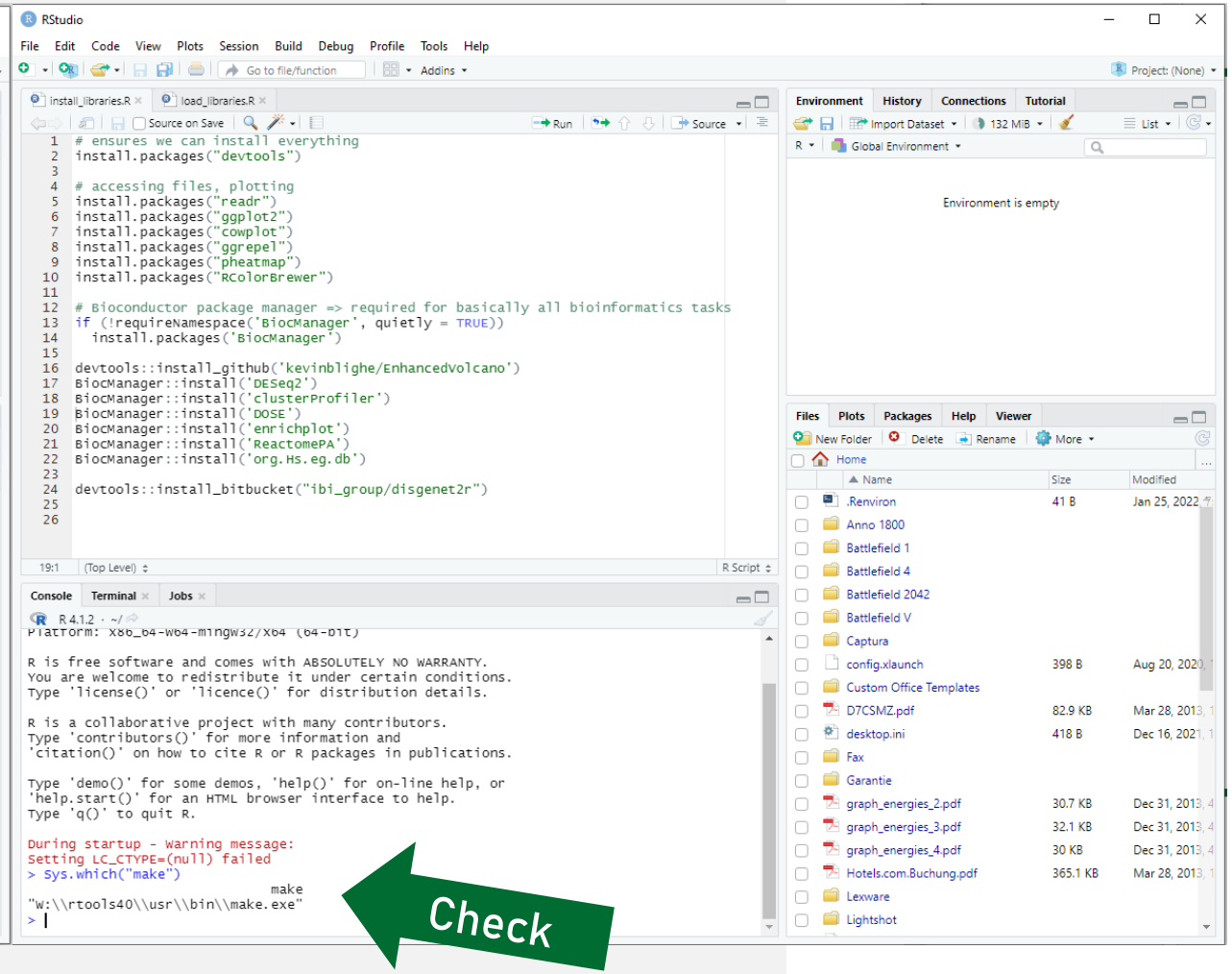
INSTALL RTOOLS



- https://cran.rstudio.com/bin/windows/Rtools/rtools40.html

MAKING RTOOLS USABLE

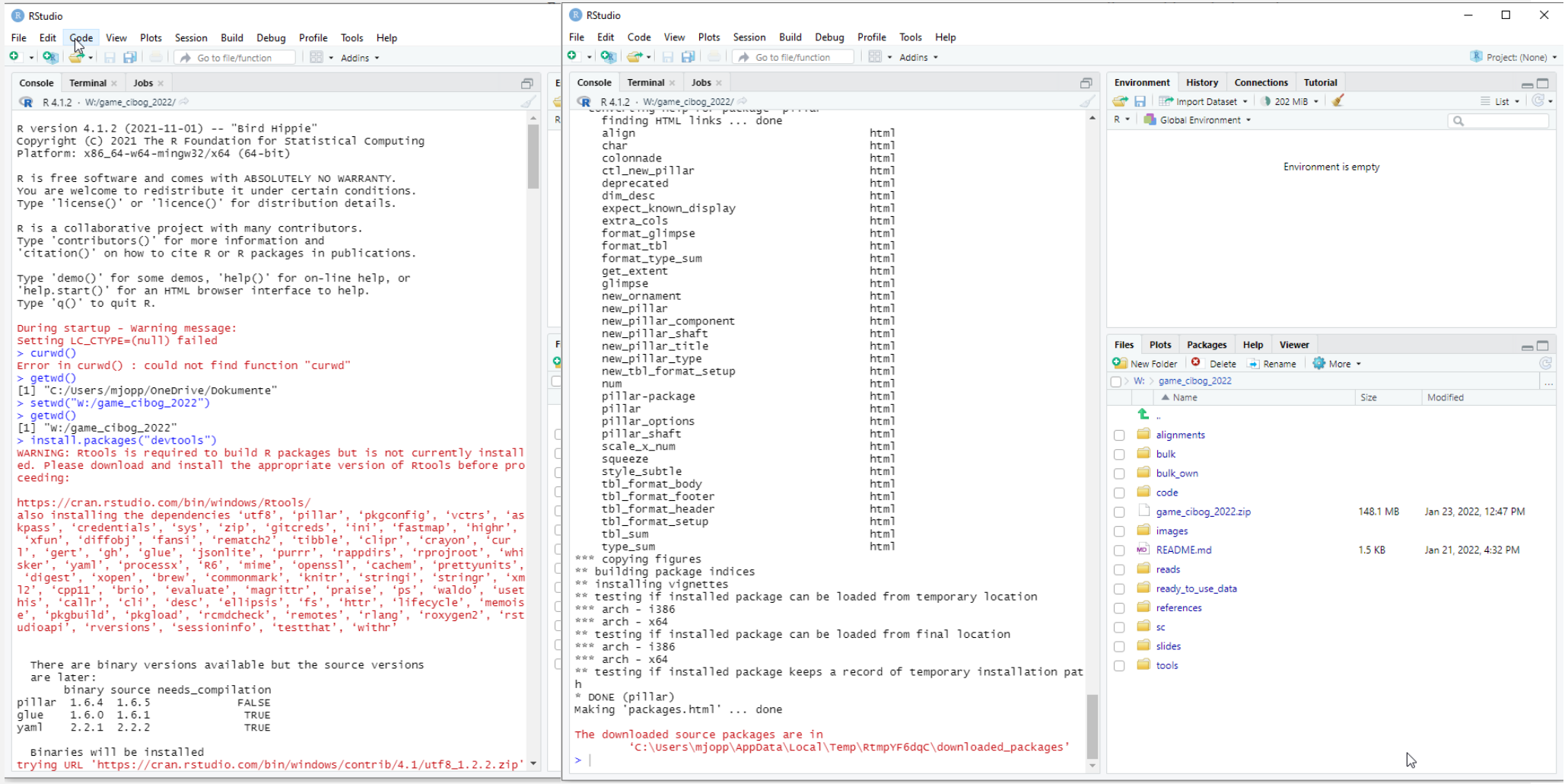




R PACKAGES

```
# ensures we can install everything
library(devtools) # => install.packages("devtools")


# accessing files, plotting
library(readr) # => install.packages("readr")
library(ggplot2) # => install.packages("ggplot2")
library(cowplot) # => install.packages("cowplot")
library(ggrepel) # => install.packages("ggrepel")
library(pheatmap) # => install.packages("pheatmap")
library(RColorBrewer) # => install.packages(" RColorBrewer ")


# Bioconductor package manager => required for basically all bioinformatics tasks
if (!requireNamespace('BiocManager', quietly = TRUE))
  install.packages('BiocManager')

library(EnhancedVolcano) # => BiocManager::install('EnhancedVolcano') or devtools::install_github('kevinblighe/EnhancedVolcano'
library(DESeq2) # => BiocManager::install('DESeq2')
library(clusterProfiler) # => BiocManager::install('clusterProfiler')
library(DOSE) # => BiocManager::install('DOSE')
library(enrichplot) # => BiocManager::install('enrichplot')
library(ReactomePA) # => BiocManager::install('ReactomePA')
library(org.Hs.eg.db) # => BiocManager::install('org.Hs.eg.db')

library(disgenet2r) # => devtools::install_bitbucket("ibi_group/disgenet2r")
```

INSTALL ALL THE LIBRARIES!



WHEN SOMETHING GOES WRONG

```
* installing *source* package 'glue' ...
** package 'glue' successfully unpacked and MD5 sums checked
ERROR: cannot remove earlier installation, is it in use?
* removing 'W:/R/R-4.1.2/library/glue'
* restoring previous 'W:/R/R-4.1.2/library/glue'
Warning in file.copy(lp, dirname(pkgdir), recursive = TRUE, copy.date = TRUE) :
  problem copying W:\R\R-4.1.2\library\00LOCK-glue\glue\libs\x64\glue.dll to W:\R\R-4.1.2
\library\glue\libs\x64\glue.dll: Permission denied

The downloaded source packages are in
        'C:\Users\mjopp\AppData\Local\Temp\RtmpW8r8CR\downloaded_packages'
Warning messages:
1: package(s) not installed when version(s) same as current; use `force = TRUE` to
  re-install: 'DOSE'
2: In install.packages(update[instlib == l, "Package"], l, repos = repos,  :
  installation of package 'glue' had non-zero exit status
>
```

```
* installing *source* package 'glue' ...
** package 'glue' successfully unpacked and MD5 sums checked
ERROR: cannot remove earlier installation, is it in use?
* removing 'W:/R/R-4.1.2/library/glue'
```

Console | Terminal | Jobs

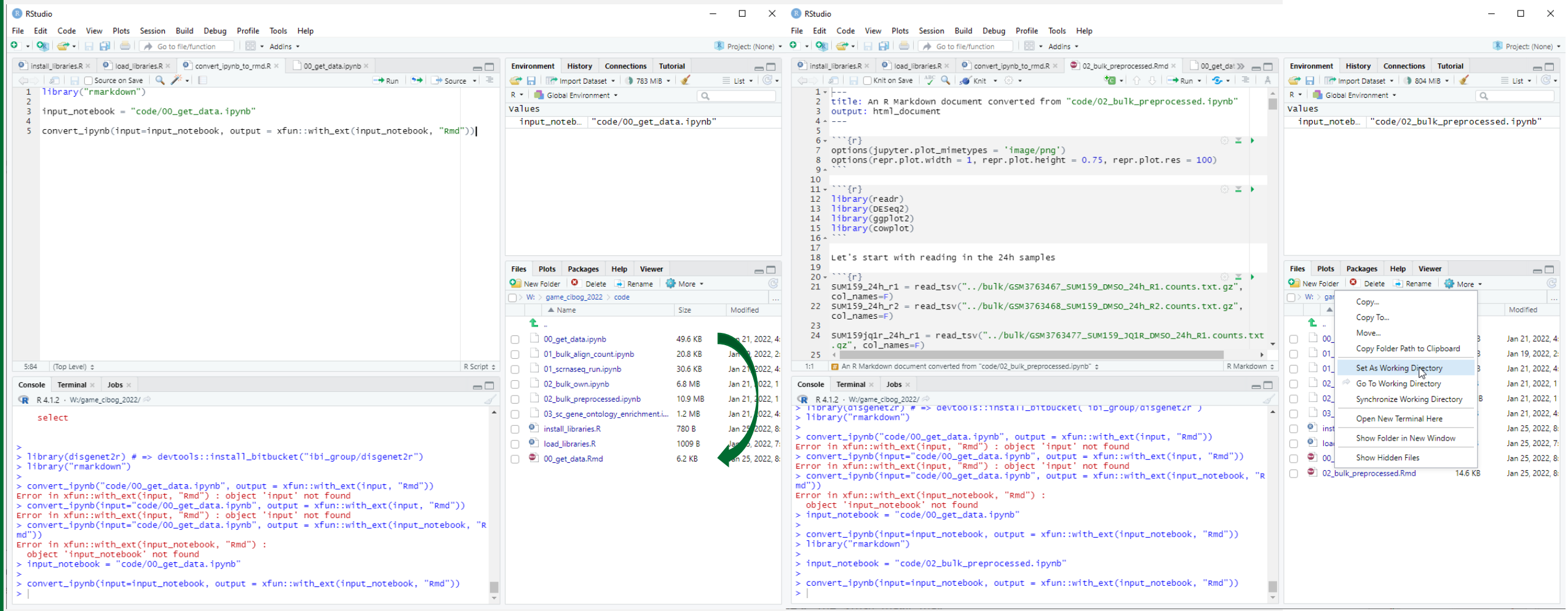
R 4.1.2 · W:/game_cibog_2022/

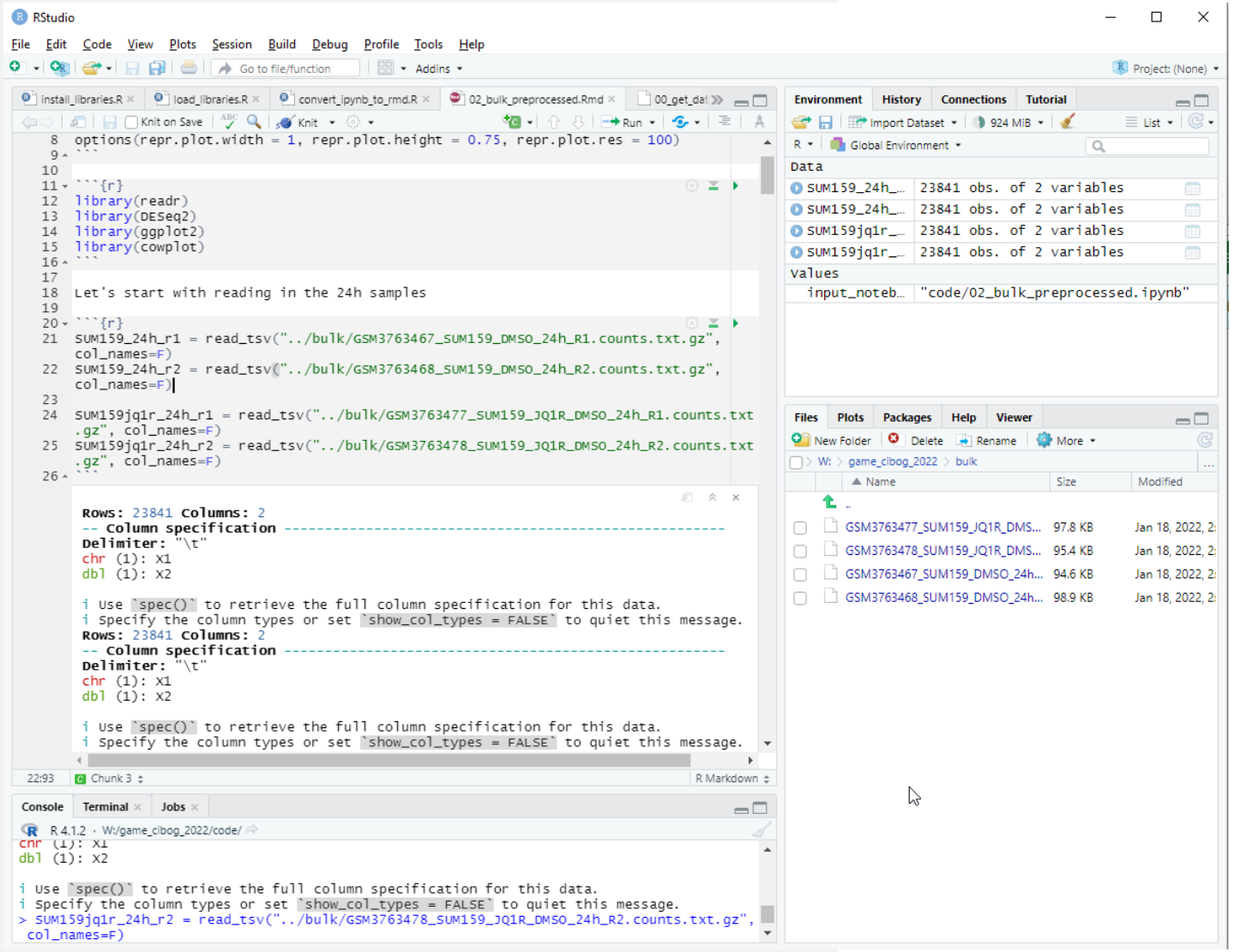
```
Content type 'application/x-gzip' length 110723 bytes (108 KB)
downloaded 108 KB

* installing *source* package 'glue' ...
** package 'glue' successfully unpacked and MD5 sums checked
ERROR: cannot remove earlier installation, is it in use?
* removing 'W:/R/R-4.1.2/library/glue'
* restoring previous 'W:/R/R-4.1.2/library/glue'
Warning in file.copy(lp, dirname(pkgdir), recursive = TRUE, copy.date = TRUE) :
  problem copying W:\R\R-4.1.2\library\00LOCK-glue\glue\libs\x64\glue.dll to W:\R\R-4.1.2
\library\glue\libs\x64\glue.dll: Permission denied

The downloaded source packages are in
        'C:\Users\mjopp\AppData\Local\Temp\RtmpW8r8CR\downloaded_packages'
Warning messages:
1: package(s) not installed when version(s) same as current; use `force = TRUE` to
  re-install: 'DOSE'
2: In install.packages(update[instlib == l, "Package"], l, repos = repos,  :
  installation of package 'glue' had non-zero exit status
> BiocManager::install('DOSE', INSTALL_opts = '--no-lock')
```

--no-lock

CONVERT IPYNB TO RMD

All notebooks expect working directory to be the 'code' directory!



RUNNING THE RMD

- Execute cells
- Check output
  - It will tell you if something did not work!