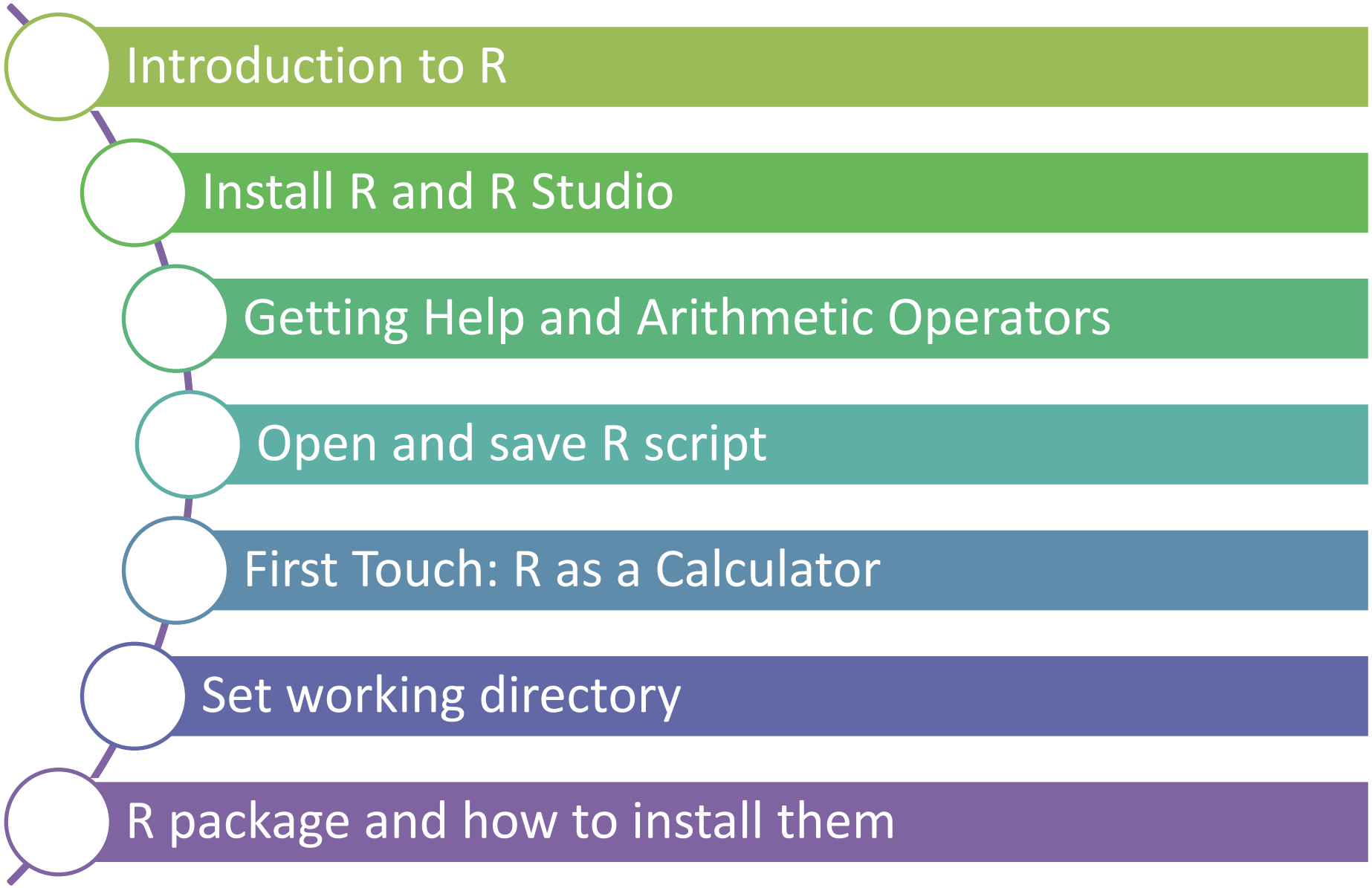
Biostatistics: Theory and Applications in R (Virtual)

Introduction to

Prof. Dr. Swapan Kumar Sarker
Department of Forestry and Environmental Science
Coordinator: BTAR Training program
Shahjalal University of Science and Technology
Sylhet, Bangladesh

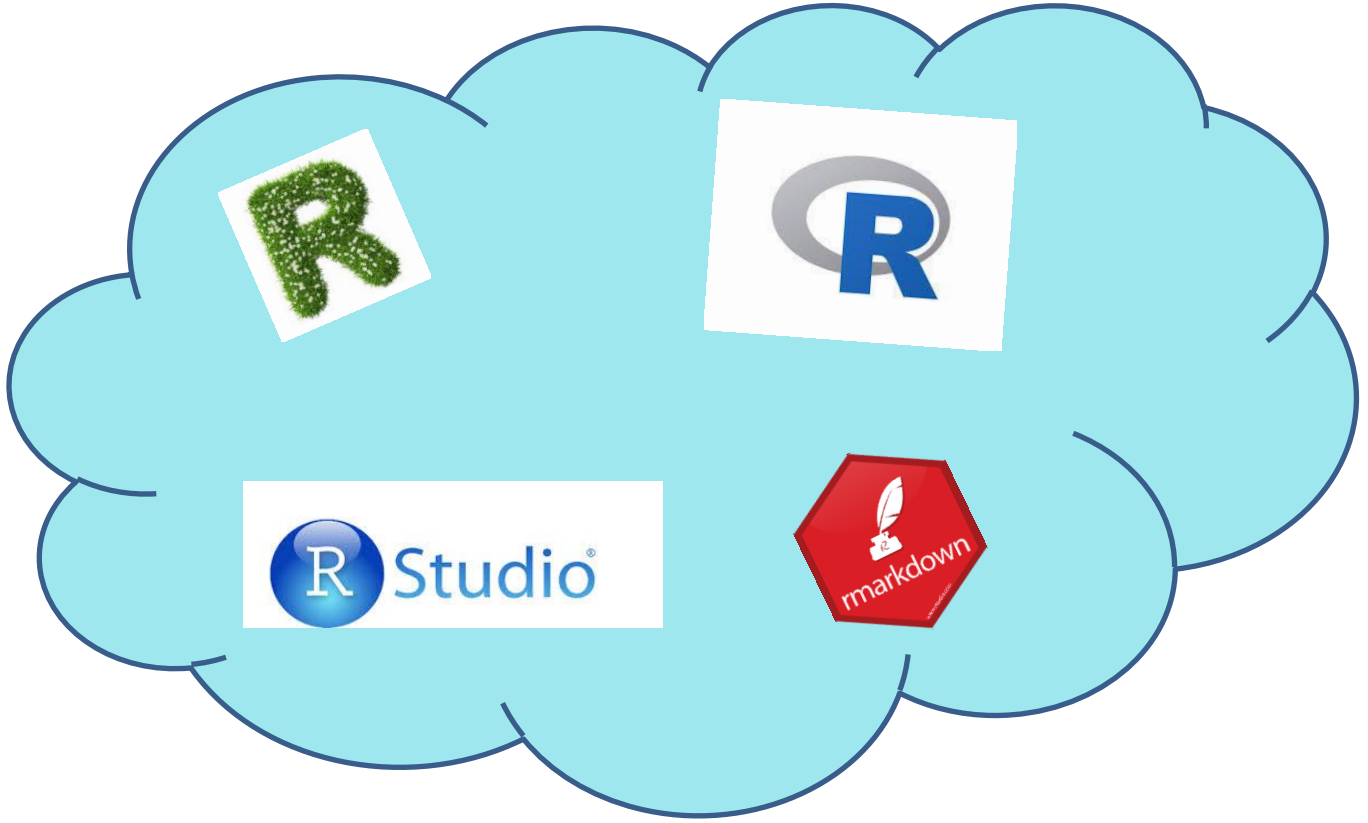
Prof. Dr. Md Abu Sayed Arfin Khan (Nobel)
Department of Forestry and Environmental Science
Co-coordinator: BTAR Training program
Shahjalal University of Science and Technology
Sylhet, Bangladesh
Khan-for@sust.edu, 01917174537

Outline of the class



What is R?

"R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, …) and graphical techniques, and is highly extensible".



Why to use R?

- ✓ a powerful tool for all sorts of statistical problems (and more)
- ✓ standard in science
- ✓ increasingly popular in ecology, economy & industry

Advantages and disadvantages of R

Advantages

- extremely powerful once you get the hang of it
- good for big data sets
- good for repetitive tasks
- recycle old code for new tasks
- no "how on earth did I do this last month?"
- lots of things you simply can't do in Excel
- free, open source

Disadvantages

- kind of scary
- steep learning curve
- not always intuitive
- slow for certain applications (but more than fast enough for what most people do with it)

How to install R?

https://cran.r-project.org/
or https://cran.rstudio.com/
or https://www.rstudio.com/products/rstudio/download/
https://posit.co/download/rstudio-desktop/

- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

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 >= 3.4.x).
old contrib	Binaries of contributed CRAN packages for outdated versions of R (for R < 3.4.x).
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.

Download R 4.2.2 for Windows (76 megabytes, 64 bit)

R-4.2.2 for Windows

Download R-4.2.2 for Windows (76 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.

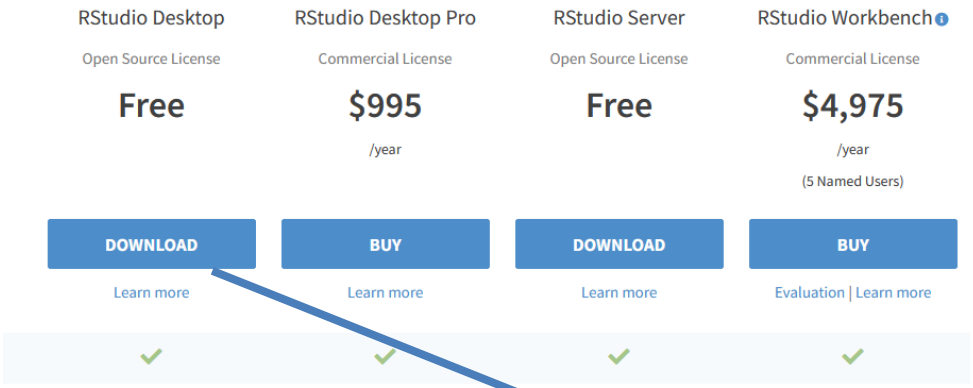
Other builds

- 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

Rstudio: Download the RStudio IDE

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

https://posit.co/download/rstudio-desktop/



RStudio®

RStudio Desktop 2022.07.2+576 - Release Notes

1. Install R. RStudio requires R 3.3.0+.

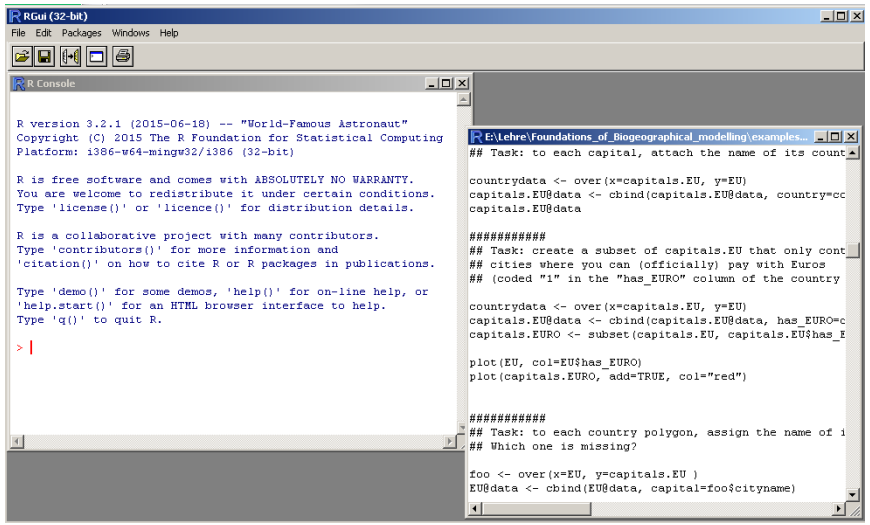
2. Download RStudio Desktop. Recommended for your system:

DOWNLOAD RSTUDIO FOR WINDOWS
2022.07.2+576 | 190.49MB

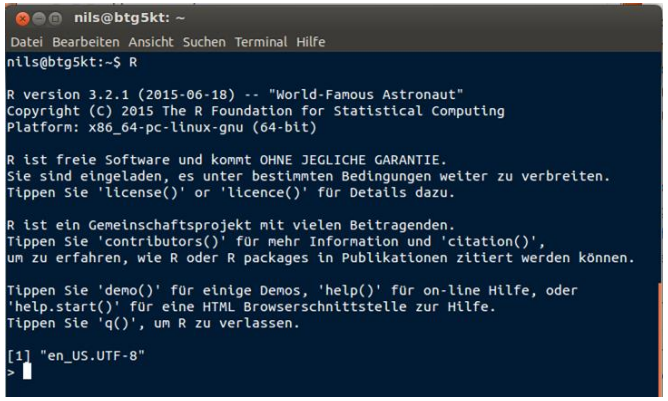
Requires Windows 10/11 (64-bit)

What does it look like?

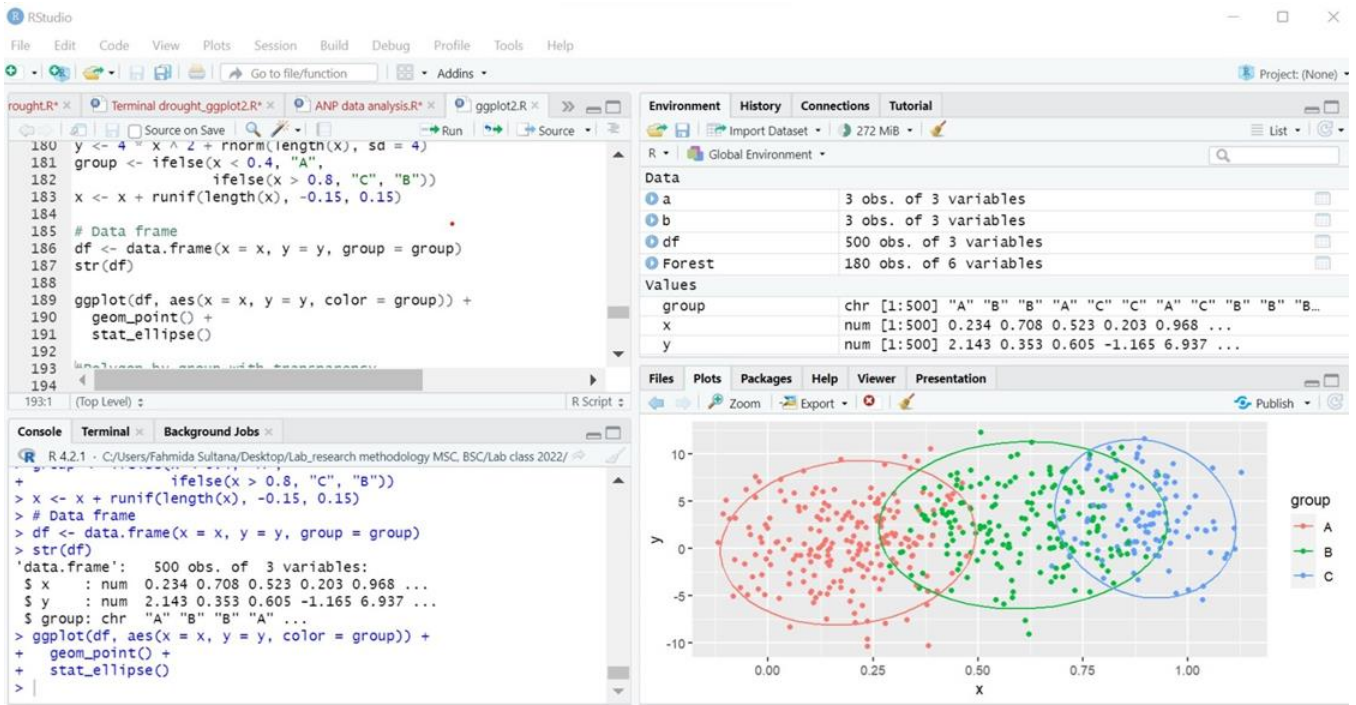
Basic built-in R GUI on Windows



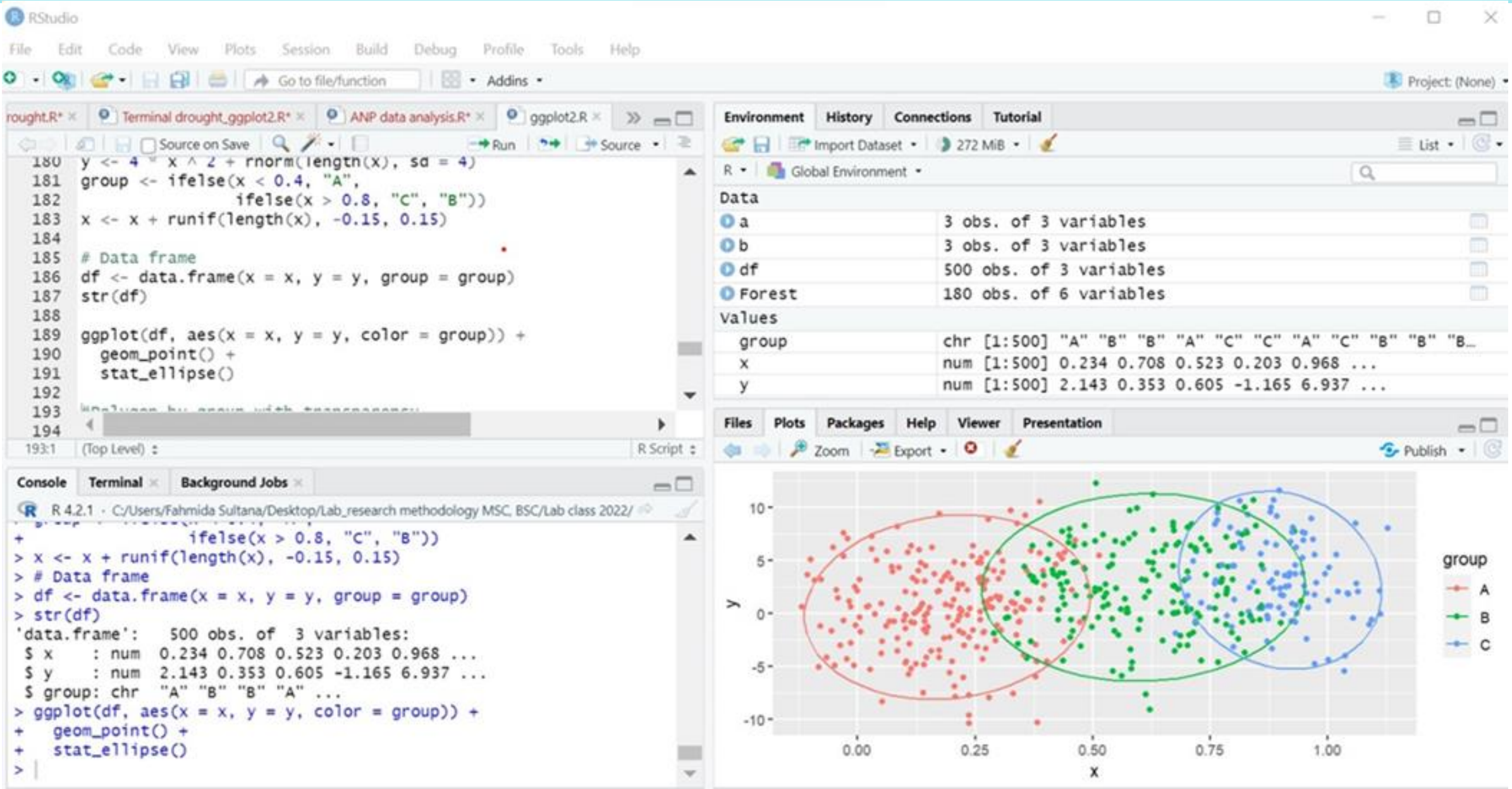
Basic R in a Linux terminal



Advanced IDE "RStudio"



RStudio


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
ggplot2.R
Source on Save
Run
Source
180 y <- 4 * x ^ 2 + rnorm(length(x), sd = 4)
181 group <- ifelse(x < 0.4, "A",
182 ifelse(x > 0.8, "C", "B"))
183 x <- x + runif(length(x), -0.15, 0.15)
184
185 # Data frame
186 df <- data.frame(x = x, y = y, group = group)
187 str(df)
188
189 ggplot(df, aes(x = x, y = y, color = group)) +
190 geom_point() +
191 stat_ellipse()
192
193
194
193:1 (Top Level) R Script
Environment History Connections Tutorial
Import Dataset 272 MiB List
R Global Environment
Data
a 3 obs. of 3 variables
b 3 obs. of 3 variables
df 500 obs. of 3 variables
Forest 180 obs. of 6 variables
Values
group chr [1:500] "A" "B" "B" "A" "C" "C" "A" "C" "B" "B" "B...
x num [1:500] 0.234 0.708 0.523 0.203 0.968 ...
y num [1:500] 2.143 0.353 0.605 -1.165 6.937 ...
Files Plots Packages Help Viewer Presentation
Zoom Export Publish
Console Terminal Background Jobs
R 4.2.1 · C:/Users/Fahmida Sultana/Desktop/Lab_research methodology MSC, BSC/Lab class 2022/
+ ifelse(x > 0.8, "C", "B"))
> x <- x + runif(length(x), -0.15, 0.15)
> # Data frame
> df <- data.frame(x = x, y = y, group = group)
> str(df)
'data.frame': 500 obs. of 3 variables:
$ x : num 0.234 0.708 0.523 0.203 0.968 ...
$ y : num 2.143 0.353 0.605 -1.165 6.937 ...
$ group: chr "A" "B" "B" "A" ...
> ggplot(df, aes(x = x, y = y, color = group)) +
+ geom_point() +
+ stat_ellipse()
>
10
5
0
-5
-10
y
0.00
0.25
0.50
0.75
1.00
x
group
A
B
C

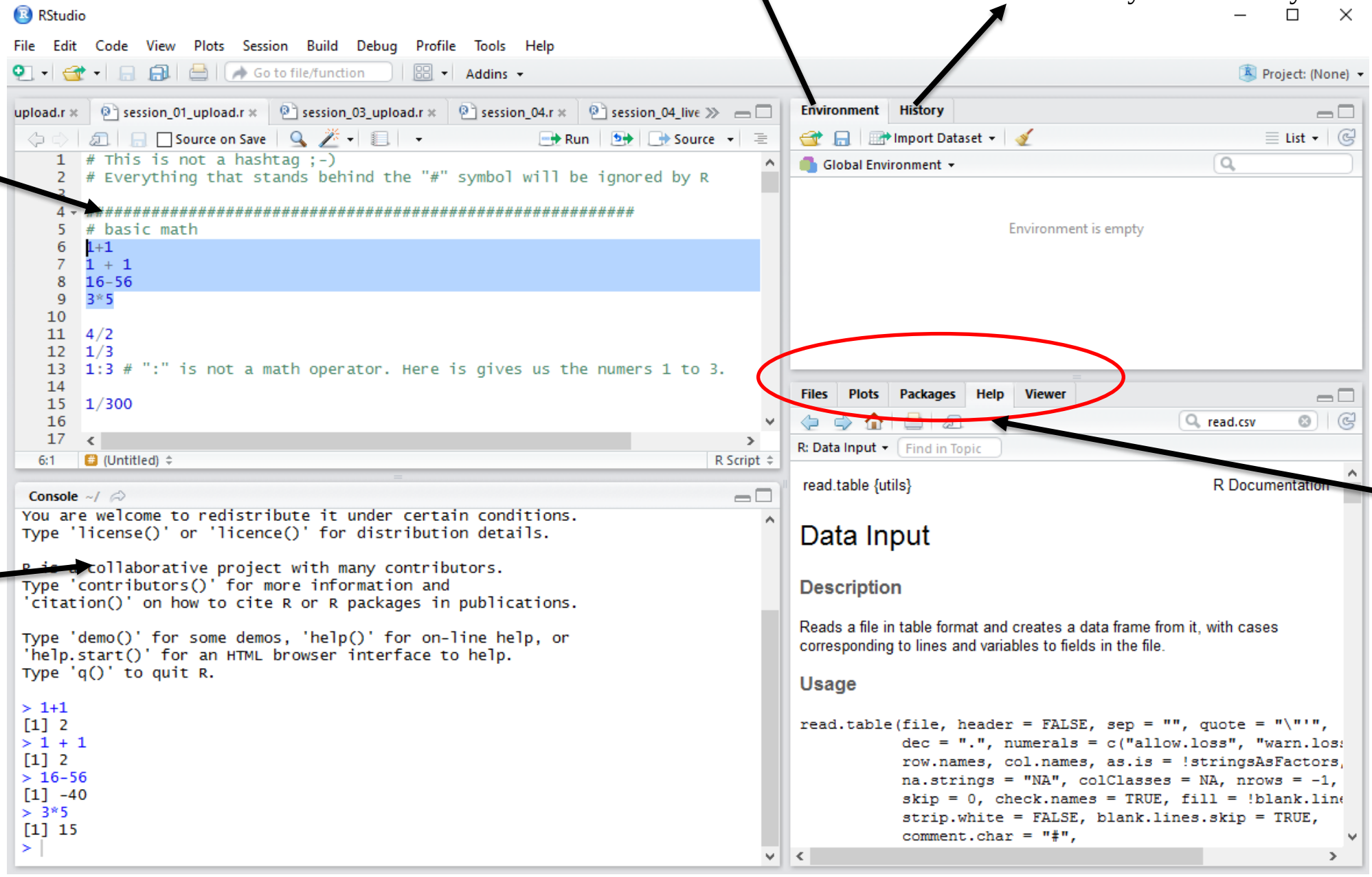
Present the windows

Workspace- tells you the variables that you`ve created

Saves the commands that you used and you can reuse them

r source editor -type comands and scripts

Console Execute script



Files
Plots
Packages
Help
Viewer

Getting Help

- Overview at r-project.org/help
- FAQs at CRAN
- Official manual "An Introduction to R"
  - as html (for online viewing)
  - as PDF (for downloading)
- Many more in different languages
- Built-in Help
  - type **help.start()** for interactive help in web browser
  - type **?functionname** for documentation of specific functions/commands (example: **?print**)

- single letters like "R" are difficult for search engines
  - a customized Google search for R is available at http://rseek.org/

Getting Help

- Overview at r-project.org/help
- FAQs at CRAN
- Official manual "An Introduction to R"
  - as html (for online viewing)
  - as PDF (for downloading)
- Many more in different languages
- Built-in Help
  - type **help.start()** for interactive help in web browser
  - type **?functionname** for documentation of specific functions/commands (example: **?print**)

- single letters like "R" are difficult for search engines
  - a customized Google search for R is available at http://rseek.org/

Arithmetic Operators

Operator	Description
+	addition
-	subtraction
*	multiplication
/	division
^ or **	exponentiation

The ^ operator raises the number to its left to the power of the number to its right: for example 3^2 is 9.

Introduction to R

Presentation

- Introduction to R
- Install R and R Studio
- Getting Help
- Arithmetic Operators

Done

R session

- Open and save R script
- First Touch: R as a Calculator
- Set working directory
- R package and how to install them

Please open RStudio

Thank you