



# *R* TUTORIALS for biologists

Weigang Qiu, Ph.D.

Department of Biological Sciences

Hunter College, City University of New York (CUNY)

Brandon Ely

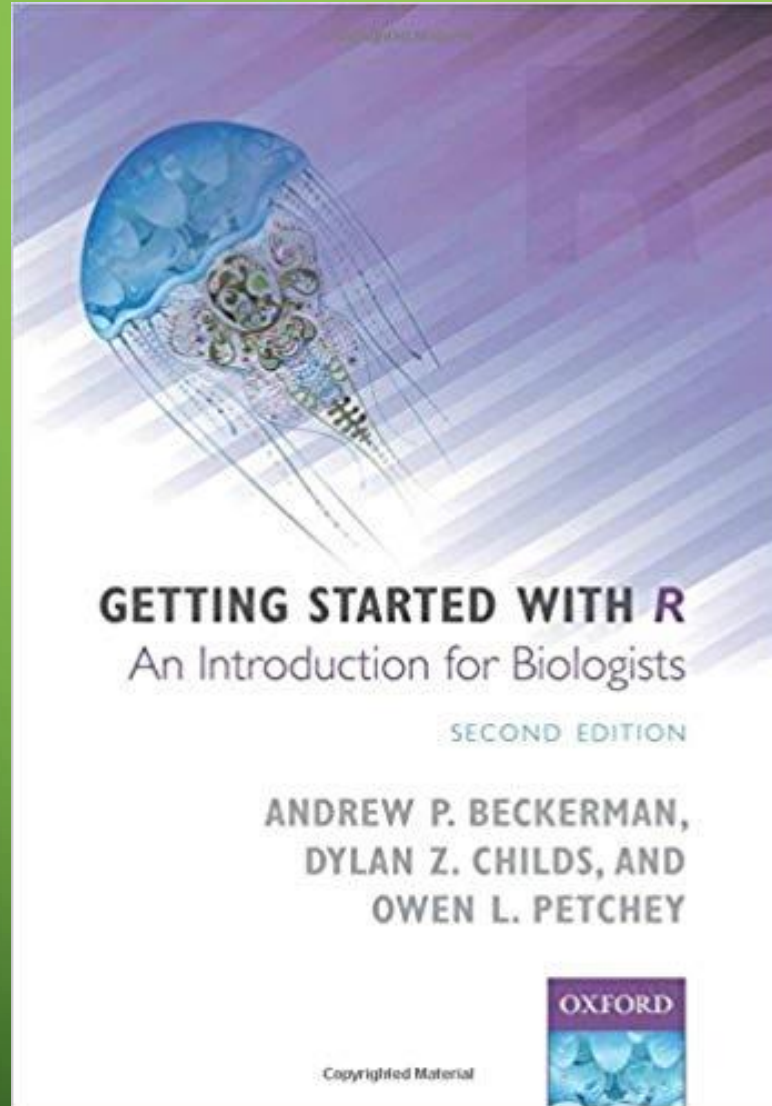
CUNY Graduate Center

August 2024

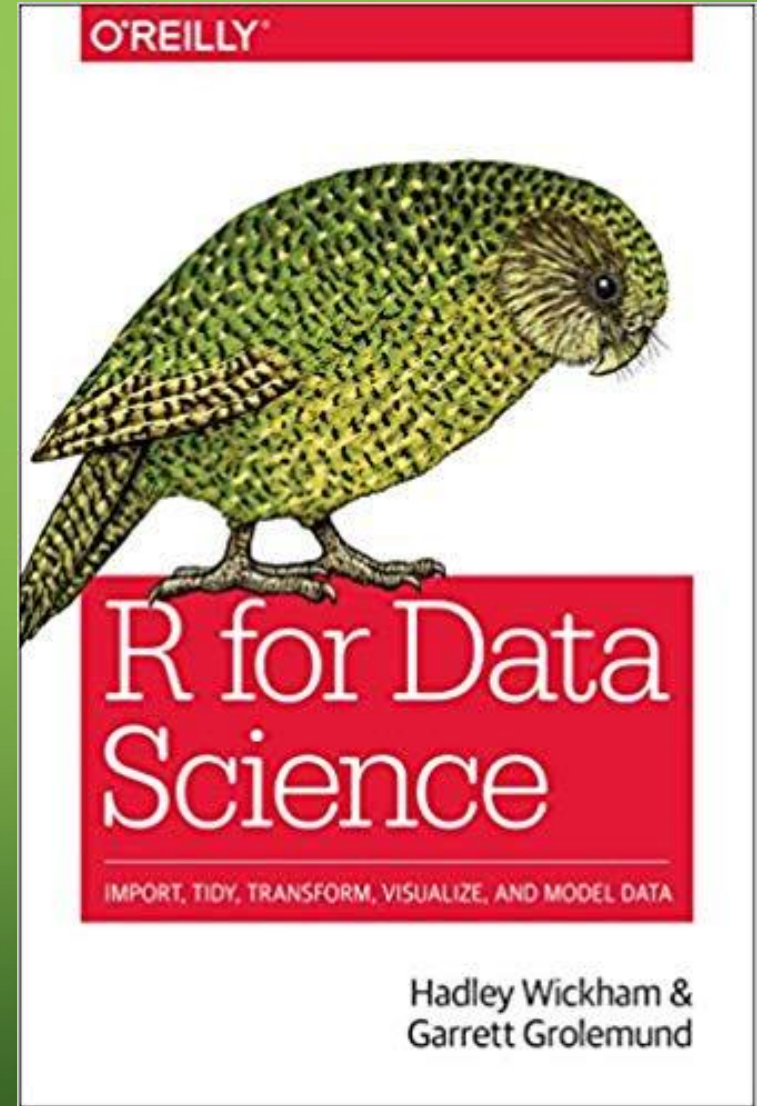
NYRaMP Informatics Workshop, DNA Learning Center @ City Tech

Workshop website: <https://wiki.genometracker.org/w/NYRaMP-Informatics-2024>

Qiu Lab wiki: [https://wiki.genometracker.org/w/Main\\_Page](https://wiki.genometracker.org/w/Main_Page)



Datasets



Online version

# INSTALLATION

- Download and install R base/RStudio:  
<https://posit.co/download/rstudio-desktop/>
- Install, load & update “tidyverse”:
  - `install.packages("tidyverse")`
  - `library("tidyverse")`
  - `tidyverse_update()`

The background is a solid green color with a subtle gradient. In the four corners, there are decorative white line-art patterns resembling circuit traces or neural network connections. These patterns consist of straight lines of varying lengths and thicknesses, ending in small white circles. The patterns are more dense in the bottom-left and top-left corners and more sparse in the top-right and bottom-right corners.

# TUTORIAL 1

## Getting Started

# Programming Terminology

- Windows/Panels

- R Console
- Script
- History/Environment
- Help/Plot/Packages

- Programming

- Variable (cAsE Sensitive!!)
- Function
- Argument
- Object (data container)

The screenshot shows the RStudio interface with the following components:

- Script Editor:** Contains the following R code:

```
1 library(tidyverse)
2 setwd("../Dropbox/Courses/ChongQing-2019/")
3 |
```
- Console:** Shows the execution of the script and subsequent sampling operations:

```
C:/Users/lai/Dropbox/Courses/ChongQing-2019/
[7] "Colin"      "Jack"
[9] "April"      "Nicolus"
[11] "Emma"
> sample(roster,1)
[1] "Jack"
> sample(roster,1)
[1] "Emma"
> sample(roster,1)
[1] "Emma"
> |
```
- Environment/History:** Shows the execution history of the script, including the creation of the 'roster' variable and the use of 'sample' function.

```
roster<-scan("roster.txt")
roster<-scan("roster.txt", what = "c...
roster
Sys.getlocale()
roster<-scan("roster.txt", what = "c...
roster
sample(roster,1)
sample(roster,1)
sample(roster,1)
```
- Package Manager:** Shows a list of installed and available packages in the 'User Library' section.

| Name                                | Description   | Version  |
|-------------------------------------|---|----------|
| <input type="checkbox"/> askpass    | Safe Password Entry for R, Git, and SSH                 | 1.1      |
| <input type="checkbox"/> assertthat | Easy Pre and Post Assertions                            | 0.2.1    |
| <input type="checkbox"/> backports  | Reimplementations of Functions Introduced Since R-3.0.0 | 1.1.4    |
| <input type="checkbox"/> base64enc  | Tools for base64 encoding                               | 0.1-3    |
| <input type="checkbox"/> BH         | Boost C++ Header Files                                  | 1.69.0-1 |
| <input type="checkbox"/> broom      | Convert Statistical Analysis Objects into Tidy Tibbles  | 0.5.2    |
| <input type="checkbox"/> callr      | Call R from R   | 3.2.0    |
| <input type="checkbox"/> cellranger | Translate Spreadsheet Cell Ranges to Rows and Columns   | 1.1.0    |

## Calculator

```
2 * 4
## [1] 8
3/8
## [1] 0.375
11.75 - 4.813
## [1] 6.937
10^2
## [1] 100
log(10)
## [1] 2.302585
log10(10)
## [1] 1
sin(2 * pi)
## [1] -2.449294e-16
7 < 10
## [1] TRUE
```

To run:

“**ctl-enter**”

## Vector

```
1:10
## [1] 1 2 3 4 5 6 7 8 9 10
```

## Function (w. arguments)

```
seq(from = 1, to = 10, by = 1)
## [1] 1 2 3 4 5 6 7 8 9 10
```

## Object (data container)

```
x <- seq(from = 1, to = 10, by = 0.5)
```

## Vector operations

```
y <- seq(from = 101, to = 110, by = 0.5)
x + y
```

To save the output to an object:

“**<-**”

## Save to script (“.R”)

```
# Amazing R. User (your name)
# 12 January, 2021
# This script is for the analysis of coffee consumption and
# burger eating

# make these packages and their associated functions
# available to use in this script
library(dplyr)
library(ggplot2)

# clear R's brain
rm(list = ls())

# Some interesting maths in R
1+1
2*4
3/8
11.75 - 4.813
10^2
log(10)
log10(10)
sin(2*pi)
x <- seq(1, 10, 0.5)
y <- seq(101, 110, 0.5)
x+y
```

Getting help: “**?seq**”

# PRACTICE #1

1. Create a variable called “**my\_first\_variable**” and assign it your last name.
2. Create a vector containing the numbers 1, 3, 0.8, and 53.3, then assign it to a variable called “**my\_numbers**”.
3. Make a vector which contain all the numbers from 6 to 12, incremented by 0.5.
4. Create and assign a vector which contain the numbers from 3 to 9. After assigning square all the numbers.
5. Create a vector of NYRaMP cohort names; apply the function “**sample()**”. Explain the output. Look up help for this function. Randomly pick a name.

```
nyramp_cohort <- c("Ahava", "Angelene", "Antonio", "Christian", "Imani", "Khizr", "Maria", "Michael", "Sugar", "Meghan")
```

6. Why didn't this code work?

```
> my_variable <- 10  
> my_variable  
Error: object 'my_variable' not found
```

7. Create a new folder (e.g., “**Desktop/r-workshop**”) & Save all commands in a file “**practice-1.R**”