



R TUTORIALS for biologists

Weigang Qiu, Ph.D.

Department of Biological Sciences

Hunter College, City University of New York (CUNY)

Brandon Ely

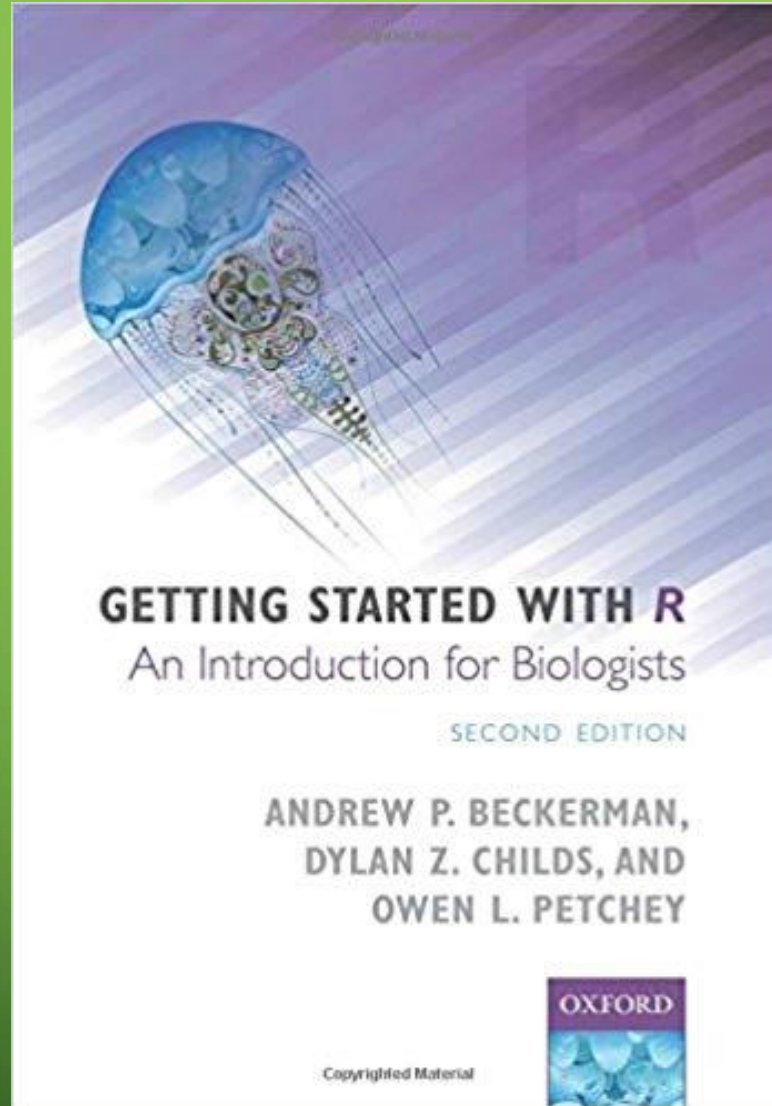
CUNY Graduate Center

August 2024

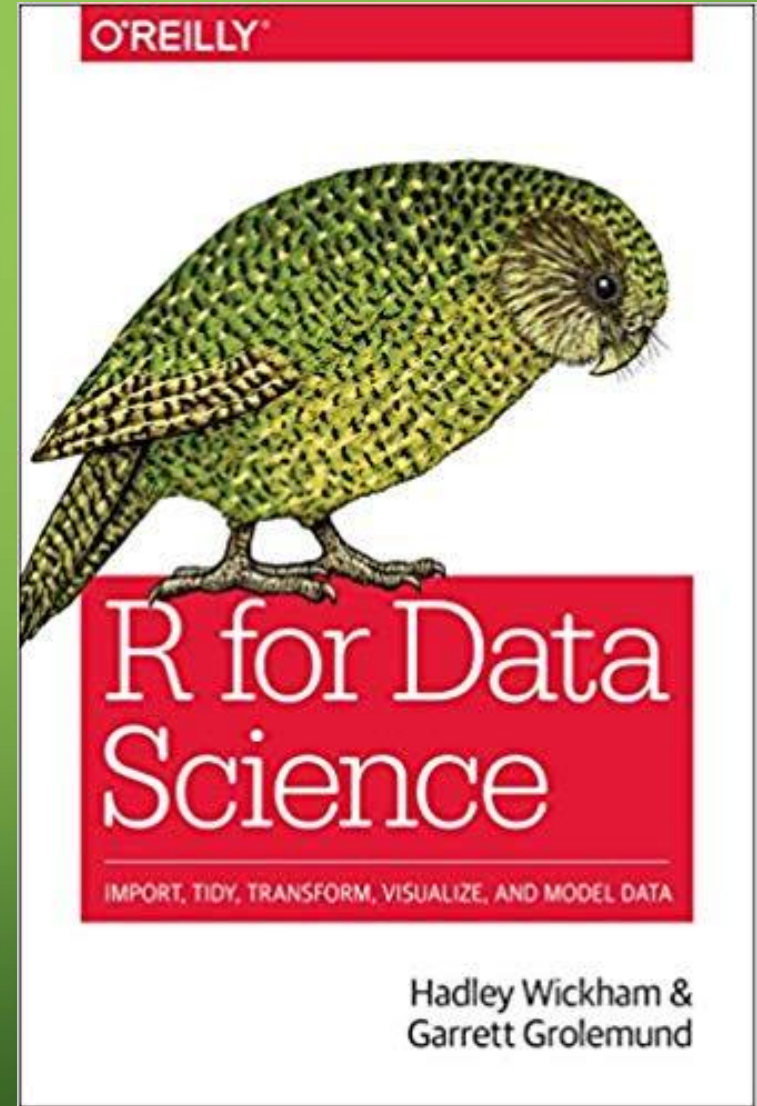
NYRaMP Informatics Workshop, DNA Learning Center @ City Tech

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

Qiu Lab wiki: https://wiki.genometracker.org/w/Main_Page



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 angles, ending in small white circles.

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 displays the RStudio environment. The top-left pane shows a script named 'roster.R' with the following code:

```
1 library(tidyverse)
2 setwd("../Dropbox/Courses/ChongQing-2019/")
3 |
```

The bottom-left pane shows the R console output:

```
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"
> |
```

The top-right pane shows the Environment pane with the following content:

```
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)
```

The bottom-right pane shows the Packages pane with a table of installed packages:

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**”

SUMMARY & TIPS

Object Function()
Unquoted Parenthesized
No space

Arguments (inputs)



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

Add spaces around operators

A character vector (**Quote the strings**)

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

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 that contains all the numbers from 6 to 12, incremented by 0.5
4. Create and assign a vector that contains the numbers from 3 to 9. After assigning, square all the numbers
5. Create a vector of DNA bases; apply the function “**sample()**”. Explain the output. Look up help for this function. Create a random DNA sequence of 1000 bases
6. Why the following codes don't work?

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

```
> data(iris)  
> glimpase(iris)  
Error in glimpase(iris) : could not find function "glimpase"
```

7. Save all commands in a file “**practice-1.R**”