DASC 3240, L05
R and Rstudio – II: R basics

2024/25 SPRING SEMESTER

MASAYUKI USHIO, ASSISTANT PROFESSOR AT OCES
USHIO@UST.HK (RM CYT-2013)

Outline (L05)

6. R basics
7. Package repositories
8. Hands-on
    - Creating objects in R
    - Data types
    - Missing data

R basics

SUMMARY OF KEY POINTS COMPARED WITH PYTHON

Difference btwn R & Python 1 | Creating an object

- You can use “=” or “<-”. Recommendation is “<-”

```
a = 5       # ok
a <- 5      # ok
b <- 3 + 5  # ok
```

Difference btwn R & Python 2 | Indexing

- Indexing start with “1”, not “0”

indexing: getting a specific element

1 2 3 4

grades = [88, 72, 93, 94]

slicing: selecting a set of elements

grades = [88, 72, 93, 94]

1 2 3 4

> grades[2:3]

Difference btwn R & Python 3 | No indent

- You can use “{}” for “for” loop and “if” statement.

```
# Create a for loop
for (i in 1:10) {
        print(i)
}

for (i in 1:10) { print(i) }

for (i in 1:10) print(i)
```

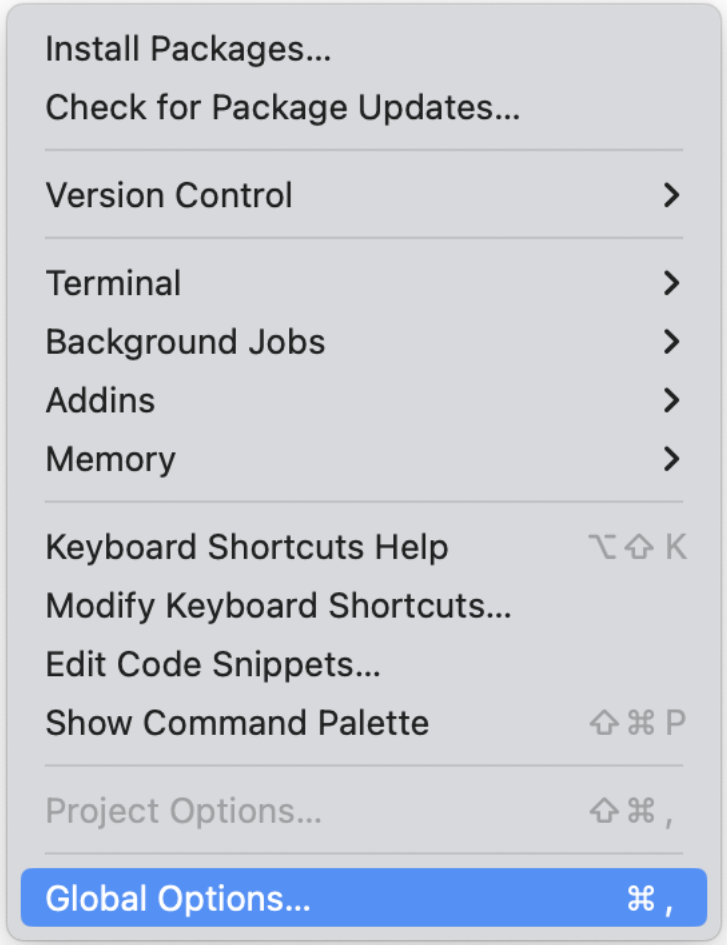
Difference btwn R & Python 3 | No indent

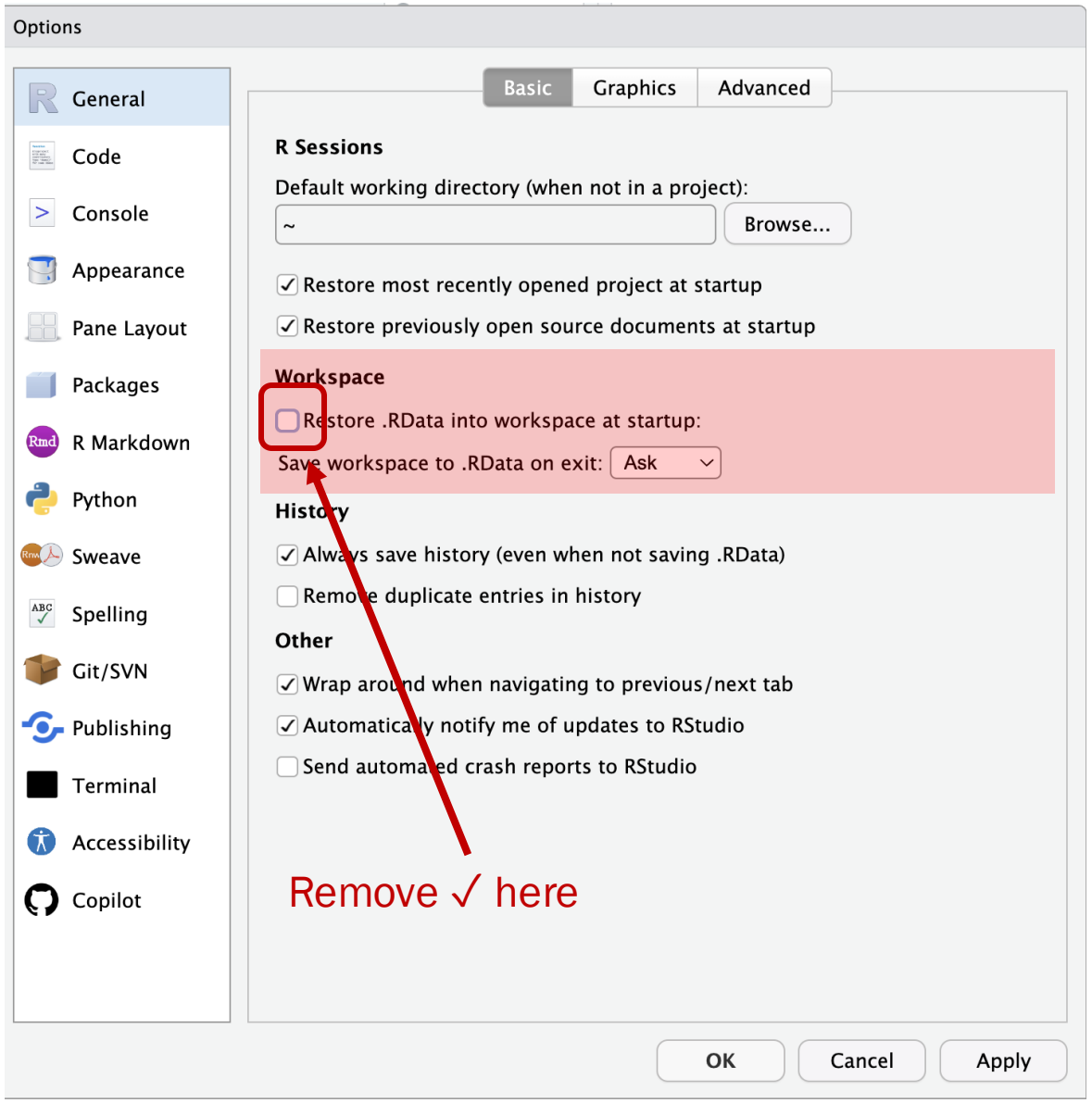
- You can use “{}” for “for” loop and “if” statement.

```
# Create an if loop
if (i == 10) {
        print(“i is 10”)
}

if (i == 10) { print(“i is 10”) }

if (i == 10) print(“i is 10”)
```

Some R tips

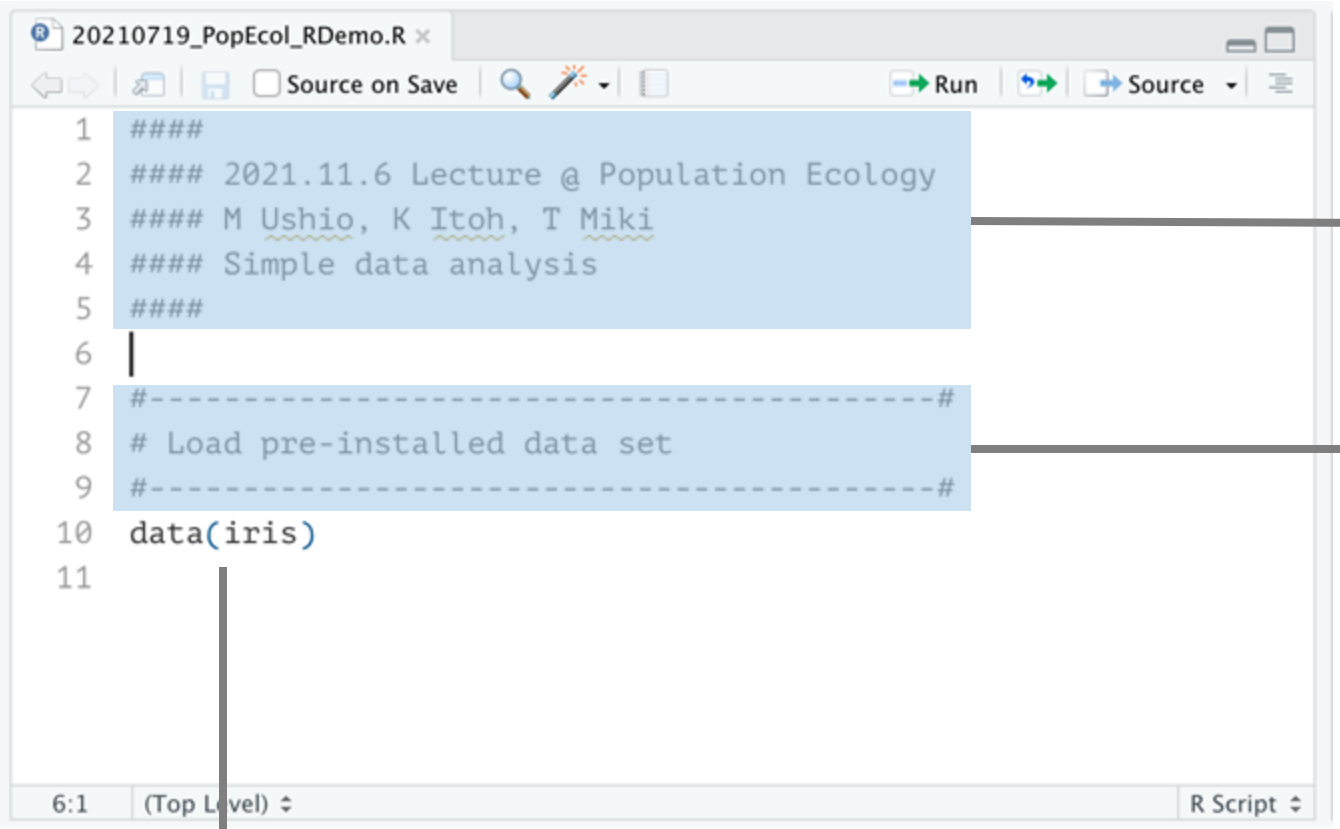
Go to “Tool” menu => “Global Options”





Some R tips

You have to make your code readable...



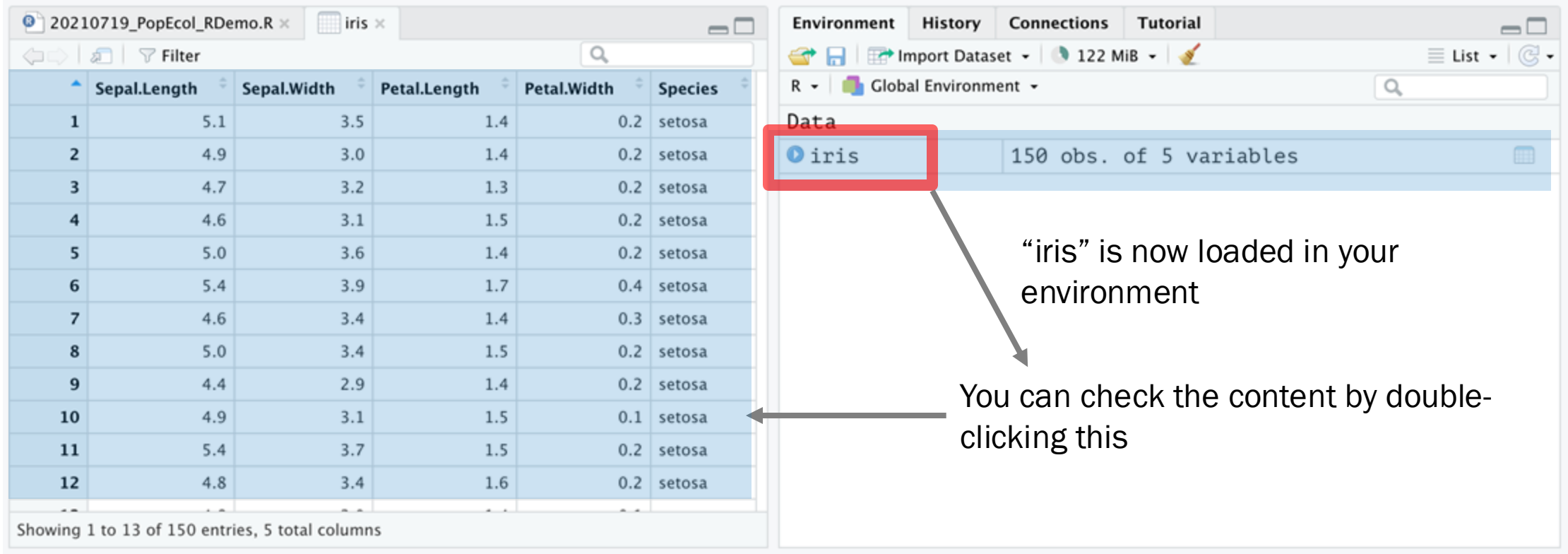
Describe code information
If you type "#", then the following characters will be recognized as a comment (not executed).
(I put four "#"s, but this is a personal preference...)

Describe what you did in this section

Caveats: Always make detailed notes and comments, even if it's a bit tedious, as I am sure that **you will forget why you wrote that code, what you did there, and what you wanted to do**

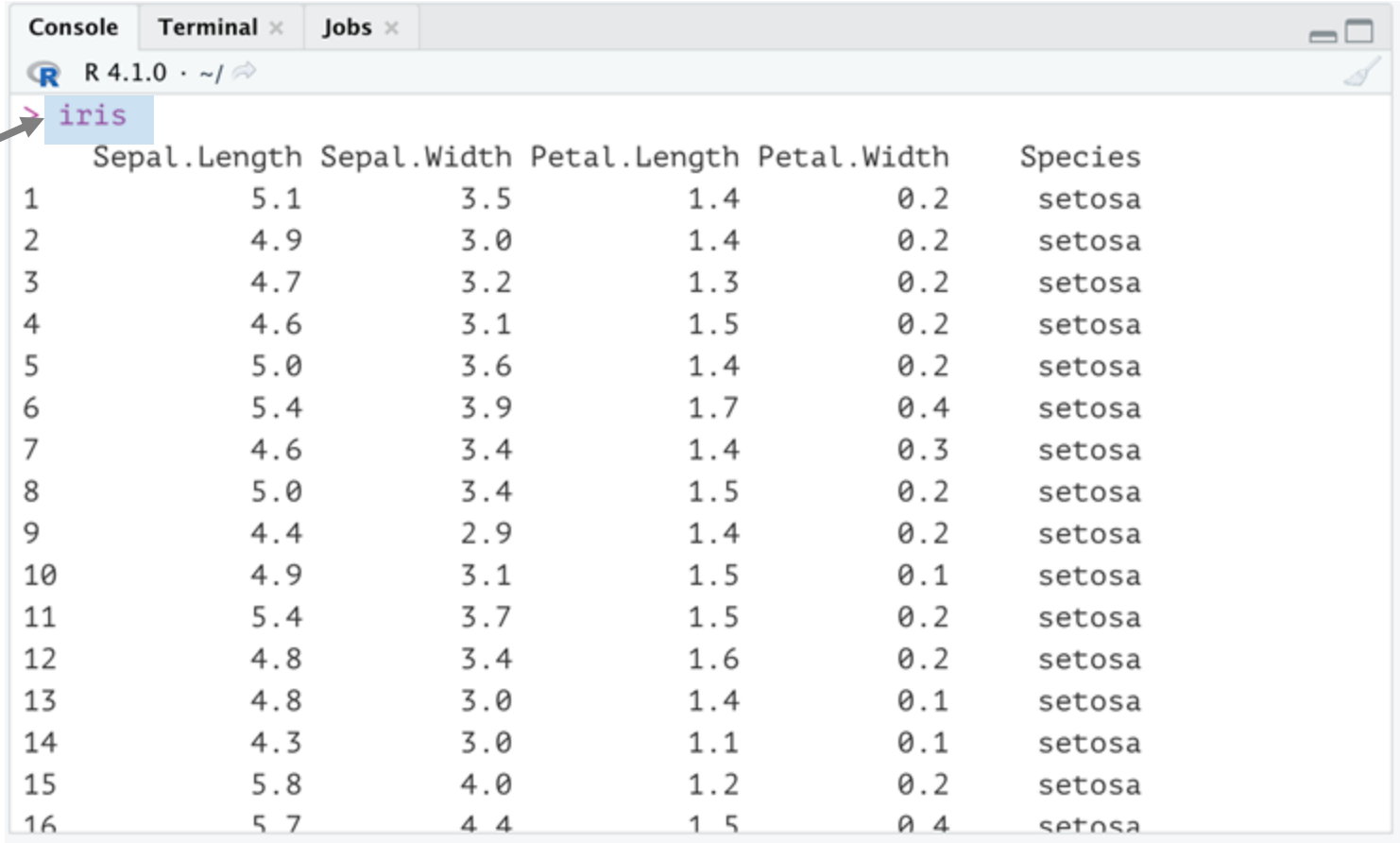
"iris" is a famous dataset included in R by default. After loading the "iris" dataset by typing "**data(iris)**", press "**Shift + Enter**" on Windows or "**Command + Enter**" on macOS to execute the command.

Some R tips

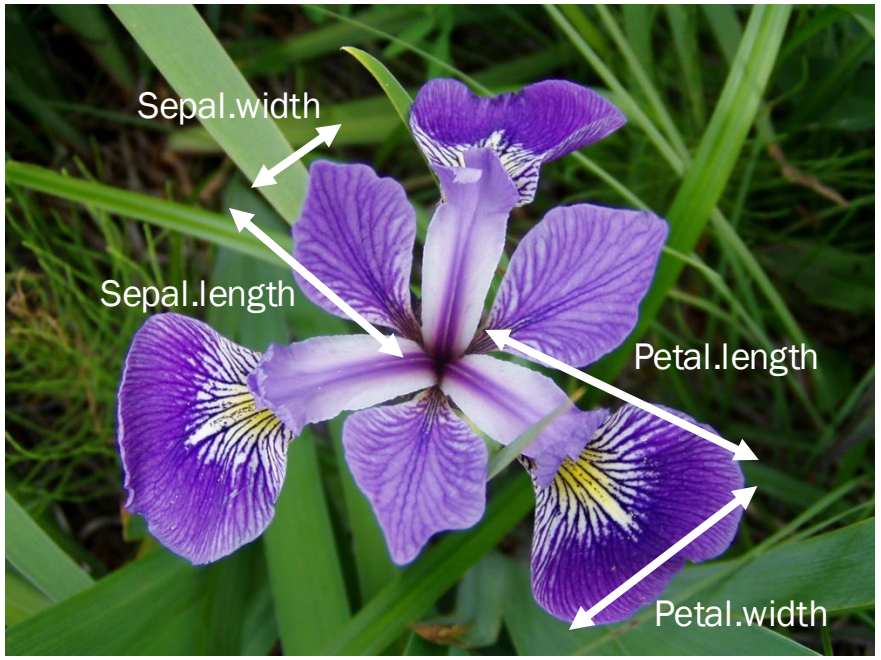

20210719_PopEcol_RDemo.R
iris
Filter
Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1 5.1 3.5 1.4 0.2 setosa
2 4.9 3.0 1.4 0.2 setosa
3 4.7 3.2 1.3 0.2 setosa
4 4.6 3.1 1.5 0.2 setosa
5 5.0 3.6 1.4 0.2 setosa
6 5.4 3.9 1.7 0.4 setosa
7 4.6 3.4 1.4 0.3 setosa
8 5.0 3.4 1.5 0.2 setosa
9 4.4 2.9 1.4 0.2 setosa
10 4.9 3.1 1.5 0.1 setosa
11 5.4 3.7 1.5 0.2 setosa
12 4.8 3.4 1.6 0.2 setosa
Showing 1 to 13 of 150 entries, 5 total columns
Environment
History
Connections
Tutorial
Import Dataset
122 MiB
List
R
Global Environment
Data
iris
150 obs. of 5 variables
"iris" is now loaded in your environment
You can check the content by double-clicking this

Some R tips

Or, you can just type “iris” in the console.



Some R tips



https://en.wikipedia.org/wiki/Iris_flower_data_set#/media/File:Iris_versicolor_3.jpg

Console Terminal Jobs

R 4.1.0 · ~/

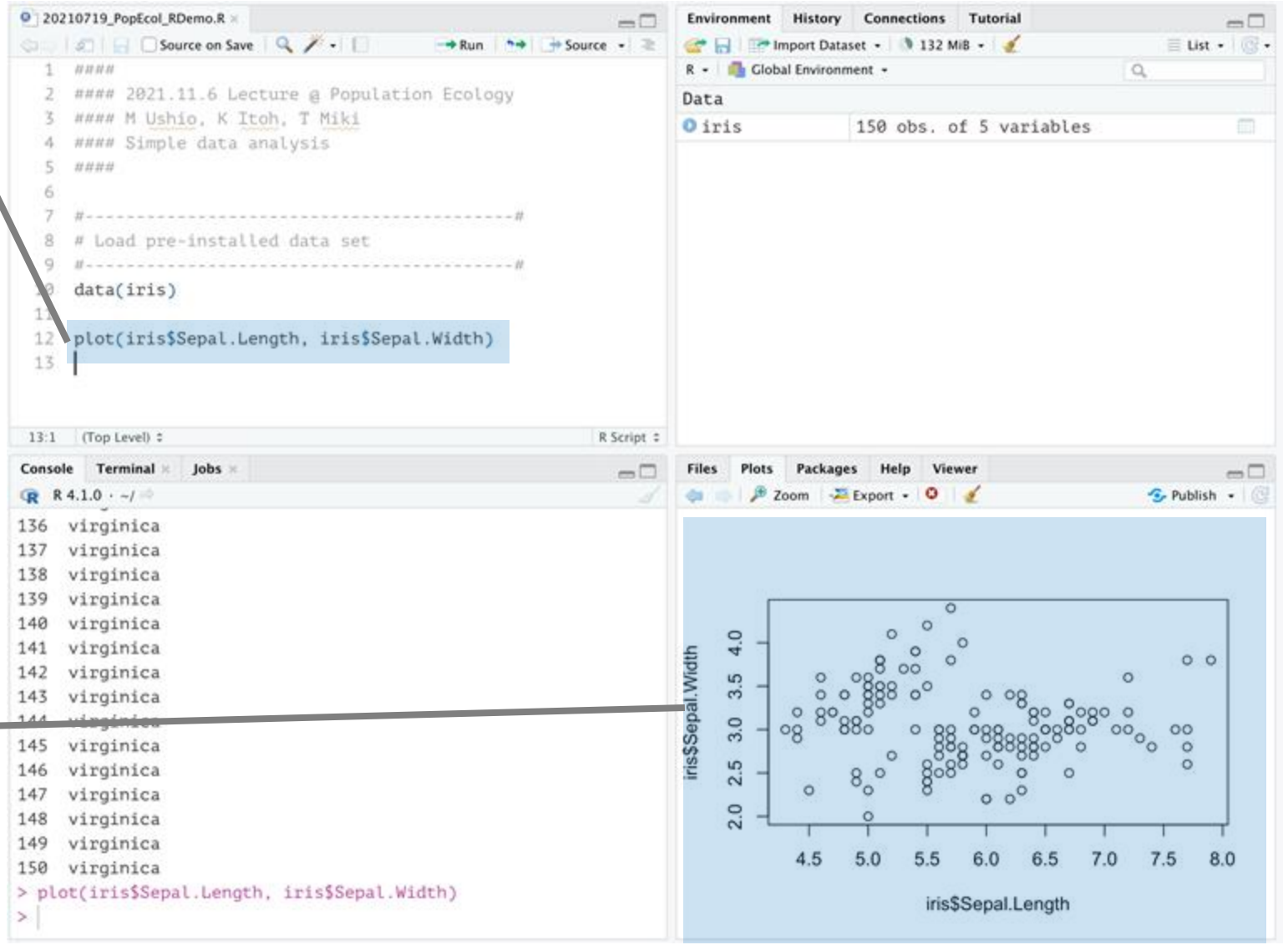
```
> iris
    Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
1            5.1         3.5          1.4         0.2     setosa
2            4.9         3.0          1.4         0.2     setosa
3            4.7         3.2          1.3         0.2     setosa
4            4.6         3.1          1.5         0.2     setosa
5            5.0         3.6          1.4         0.2     setosa
6            5.4         3.9          1.7         0.4     setosa
7            4.6         3.4          1.4         0.3     setosa
8            5.0         3.4          1.5         0.2     setosa
9            4.4         2.9          1.4         0.2     setosa
10           4.9         3.1          1.5         0.1     setosa
11           5.4         3.7          1.5         0.2     setosa
12           4.8         3.4          1.6         0.2     setosa
13           4.8         3.0          1.4         0.1     setosa
14           4.3         3.0          1.1         0.1     setosa
15           5.8         4.0          1.2         0.2     setosa
16           5.7         4.4          1.5         0.4     setosa
```

Some R tips

Visualize the relationship between variables using plot() function.

- ”$“ after “iris” is used to access a variable in the data frame.
- Visualize the relationship between “iris$Sepal.Length” and “iris$Sepal.Width"

Select the command, press “Shift+Enter” (Win) or “Command+Enter” (Mac). Then, the command will be executed and a plot will be shown at right-bottom.



Package repositories

Package repositories

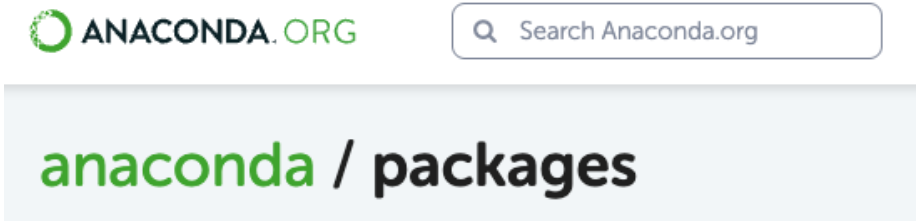
Python

- PyPi (https://pypi.org/)
- Anaconda (https://anaconda.org/)

R

- CRAN (https://cran.r-project.org/)
- Bioconductor (https://www.bioconductor.org/)

The Comprehensive R Archive Network



Package repositories

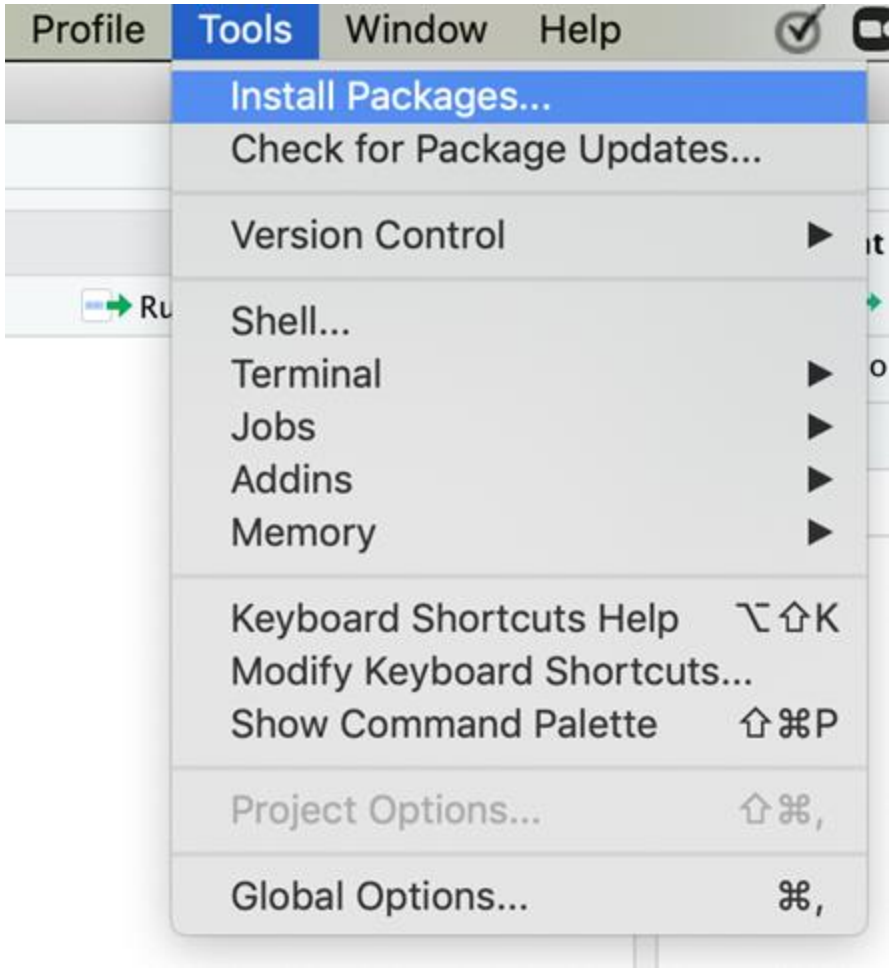
R packages are registered in **The Comprehensive R Archive Network (CRAN)**
https://cran.r-project.org/

Available CRAN Packages By Name

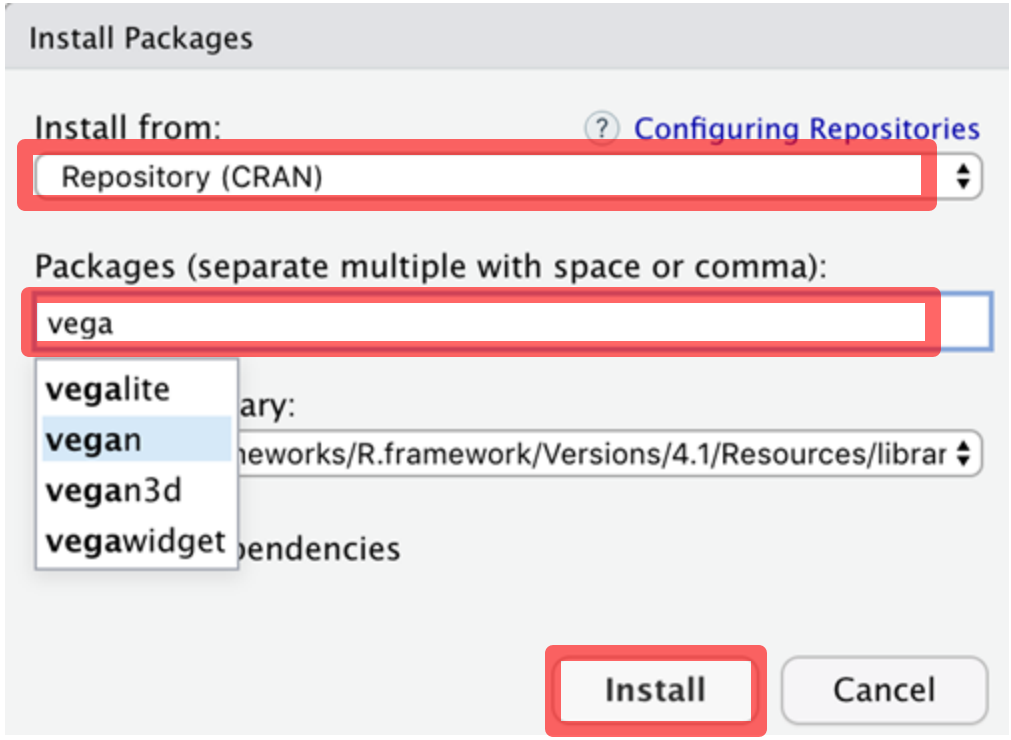
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
aaSEA	Amino Acid Substitution Effect Analyser
AATtools	Reliability and Scoring Routines for the Approach-Avoidance Task
ABACUS	Apps Based Activities for Communicating and Understanding Statistics
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcADM	Fit Accumulated Damage Models and Estimate Reliability using ABC
ABCanalysis	Computed ABC Analysis
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2

Package repositories



- Select “Tools” in the RStudio menu bar → Select “Install Packages”
- Select “CRAN” in the “Install from” tab and input a package name you want to install
- Click “Install”



Package repositories

The Comprehensive R Archive Network

```
install.packages(XXXXX)
```

```
If (!require("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
BiocManager::install("XXXXX", version ="XXX")
```

```
install.packages("remotes")
remotes::install_github("XXXXX")
```

```
install.packages("devtools")
devtools::install_github("XXXXX")
```

Package repositories

:: (specify namespace)

- In R, "::" is used to specify "namespace". In each namespace, many functions are defined. If you specify namespace, you can avoid conflicts (e.g., different packages define functions with the same name).
- "xxxx::yyyy" means that you will use a function "yyyy" defined by a package "xxxx"

```
install.packages("remotes")
remotes::install_github("XXXXX")
```

```
install.packages("devtools")
devtools::install_github("XXXXX")
```

Course materials

Presentation slides

me

you

R Studio®

R codes
Other text resources

rmarkdown

Prepared with
Rmarkdown

upload

upload

jupyter

GitHub

download is
not necessary

download

https://github.com/ong8181/DASC3240

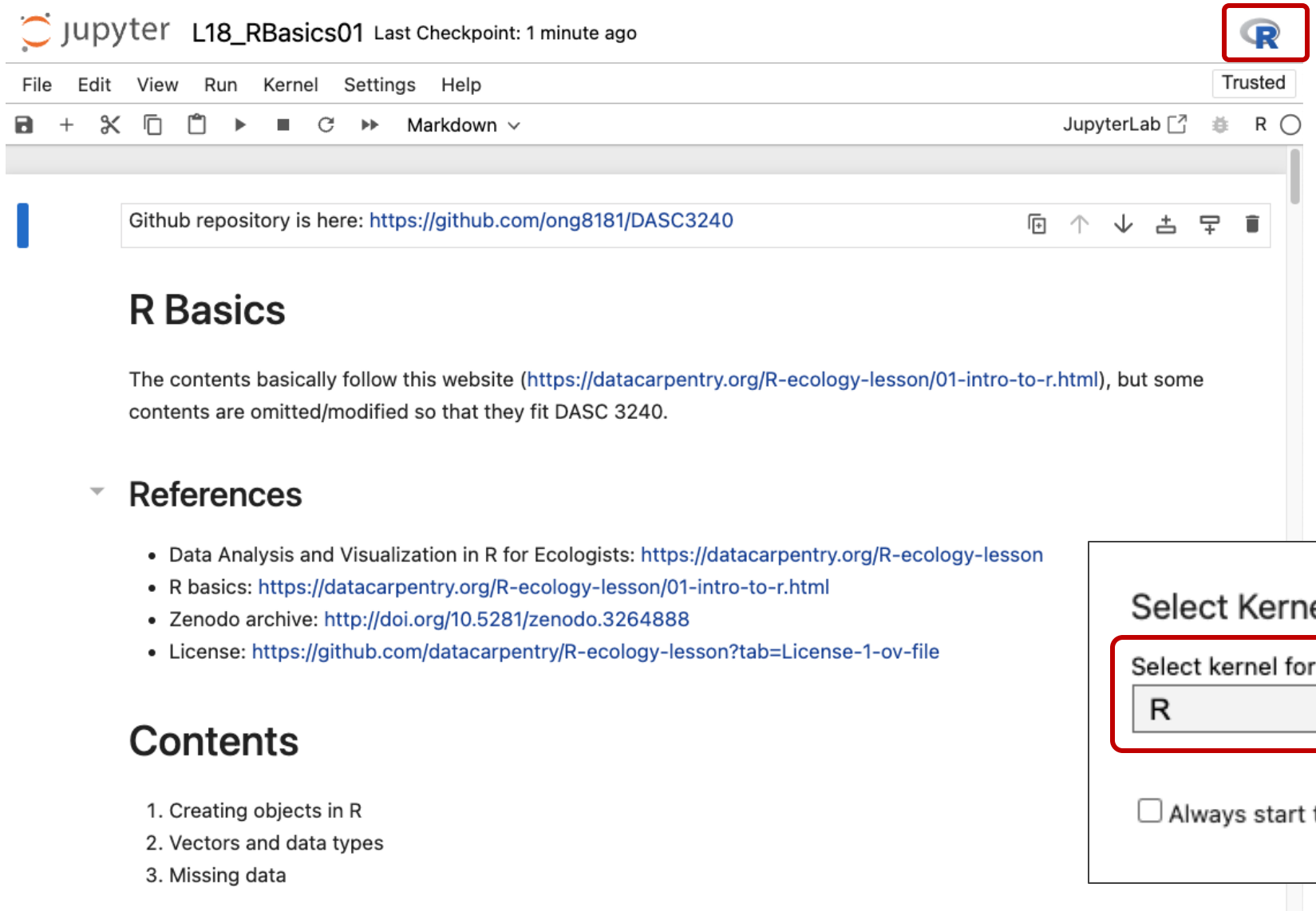
Option 1:
If you do not have R and RStudio in your PC, you can execute the code using Google Colab.

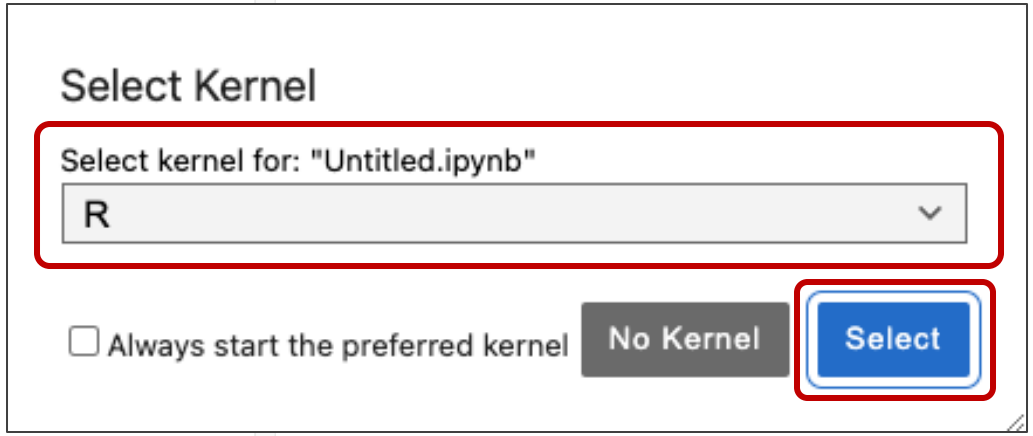
colab

Option 2:
If you have R and RStudio in your PC, you can download the code and open them using RStudio.

If you have a trouble(s), you can utilize a face-to-face assistant opportunity in the tutorial.

Jupyter notebook





Jupyter notebook

- If you cannot change the kernal (language), please try the following in R console.
  - Launch R or Rstudio
  - In R console, type

```
install.packages(c('repr', 'IRdisplay', 'evaluate', 'crayon', 'pbdZMQ', 'devtools', 'uuid', 'digest'))
install.packages('IRkernel')
IRkernel::installspec()
```

Data Carpentry

The contents of this lecture generally follow the tutorial below.

Data Analysis and Visualization in R for Ecologists

DATA CARPENTRY

BUILDING COMMUNITIES TEACHING UNIVERSAL DATA LITERACY

https://datacarpentry.org/
https://datacarpentry.github.io/R-ecology-lesson/

Licensed under CC-BY 4.0