# **Introduction to Biostatistics Class Guidance**

Hiroyoshi Iwata, hiroiwata@g.ecc.u-tokyo.ac.jp

2021/4/9

# **Contents of the lecture**

In this course, students will learn the basics of statistics by using R and RStudio.

The content covered in the lecture is as follows.

- 1. visualizing data with R (April 9)
- 2. regression and analysis of variance (April 16)
- 3. principal component analysis, multidimensional scaling constructs (April 23)
- 4. hierarchical and non-hierarchical cluster analysis (April 30)

# Lecture format

## Please download required files before the class

Each lecture is taught in a hands-on format using text (PDF, HTML, Google Docs), R code (R), and data (CSV) files. All files should be available for download from the AgriBio website and ITC-LMS at least one day before the lecture. Please make sure to download the book before the lecture.

#### There will be quizzes sometimes in the lecture

In order to add a kind of game-like play in the lecture, we will have quizzes in the lecture. Let's do a little practice here.

- 1. First, go to https://www.menti.com/95twm188to
- 2. Next, enter your nickname and click on "Join quiz".



3. Then it goes into a standby state. After a while, the quiz will start. I wish you all the best in your endeavors.



Please note that these quizzes are just a kind of game and do not affect your grades. Have fun and work on it.

# Grades

Grades will be based on attendance (which will be taken using the ITC-LMS) and reports (which will be submitted using the ITC-LMS).

# **Installing R and RStudio**

#### In the case of Windows

#### Installing R

- 1. First, go to the CRAN site https://cran.ism.ac.jp/.
- 2. Click "Download R for Windows".



3. Select "base".



4. Click "Download R 4.0.5 for Windows". The latest version will vary depending on the time.



5. When you download, please "save" instead of "run".

6. Click on the downloaded and file to install R. Please be sure to "Run as an administrator" (install with administrative privileges).



7. You do not need to change any of the installation options. You can proceed with the installation with the default settings.

## **Installing R Studio**

- 1. Go to the RStudio download site: https://rstudio.com/products/rstudio/.
- 2. Next, click "RStudio Desktop".



3. Click the "DOWNLOAD RSTUDIO DESKTOP".

$\leftarrow \ \rightarrow \ \bigcirc$	6	https://rstudio.com/products/rstudio/	□ ☆ ☆	± l.	. Ŀ	••••
		diagnose and fix errors quickly • Extensive package development tools				~
	Support	Community forums only	<ul> <li>Priority Email Support</li> <li>8 hour response during business hours (ET)</li> </ul>			ļ
	License	AGPL v3	RStudio License Agreement			
	Pricing	Free	\$995/year			
		DOWNLOAD RSTUDIO DESKTOP	DOWNLOAD FREE RSTUDIO DESKTOP PRO TRIAL			
			Purchase   Contact Sales			
https://rstudio.com/r	anducts/retudio/	(download)				

4. Click the "DOWNLOAD" button on the left side (free version of RStudio Desktop).



5. When you download, please "save" instead of "run".

6. Click on the downloaded file to install RStudio. Please be sure to "Run as an administrator" (install with administrative privileges).



7. You do not need to change any of the installation options. You can proceed with the installation with the default settings.

## In the case of Mac OS

#### **Installing R**

1. First, go to the CRAN site https://cran.ism.ac.jp/.

2. Click "Download R for (MAC) OS X".



3. Click "R-4.0.5.pkg". The latest version will vary depending on the time.



4. Click on the downloaded file to install R.



5. You do not need to change any of the installation options. You can proceed with the installation with the default settings.

# Installing R Studio

- 1. Go to the RStudio download site: https://rstudio.com/products/rstudio/.
- 2. Next, click "RStudio Desktop".



3. Click the "DOWNLOAD RSTUDIO DESKTOP".

R Stu	dio Desktop		
	Open Source Edition	RStudio Desktop Pro	
Overview	Access RStudio locally     Syntax highlighting, code completion, and smart indentation     Execute R code directly from the source editor     Quickly jump to function definitions     Easily manage multiple working directories using projects     Integrated R help and documentation     Interactive debugger to diagnose and fix errors quickly     Extensive package development tools.	All of the features of open source; plus: • A commercial license for organizations not able to use AGPL software • Access to priority support • RStudio Professional Drivers	
Support	Community forums only	Priority Email Support     8 hour response during business hours (ET)	
License	AGPL v3	RStudio License Agreement	
Pricing	Free	\$995/year	
	DOWINLOAD RETUDIO DESKTOP	DOWNLOAD FREE RSTUDIO DESKTOP	

		i rstudio.com	Ċ	۵
	RStudio Desktop	RStudio Desktop	RStudio Server	RStudio Server Pro
	Free	\$995 /year	Free	\$4,975 /year (5 Named Users)
	DOWNLOAD	BUY	DOWNLOAD	BUY
	Learn more	Learn more	Learn more	Evaluation   Learn more
Integrated Tools for R	×	~	~	~
Priority Support		~		~
Access via Web Browser			×	~
Enterprise Security				~
Project Sharing				~
Manage Multiple R Sessions & Versions				~
Admin Dashboard				~
Load Balancing				~
Auditing and Monitoring				~

4. Click the "DOWNLOAD" button on the left side (free version of RStudio Desktop).

5. When you click on the downloaded file, a window will be displayed as follows. All that's left to do is to drag RStudio and drop it on top of Applications.

!

# Trying to use R and RStudio

# Make sure that R and RStudio are working

From here, I will explain how to use R and RStudio for Windows and Mac OS users together, because RStudio has the same interface in Windows and Mac OS versions.

Let's start by typing the following into the "Console" on the left of the window.

## 3 + 4 \* 2

## [1] 11



The calculation is performed and the result is shown. I will explain how to use R in detail in next week's lecture.

#### Preparing to use R and RStudio for the lecture

#### Install R packages.

R can perform a variety of statistical analyses even as it is, but if you install an "extended set" called a package, the number of available analytical methods can be largely expanded. This section explains how to install a package. Please note that your computer must be connected to the Internet in order to perform the installation of a package.

1. Select "Install Packages" from the Tools menu in RStudio.



2. Enter the name of the package you want to install in the dialog box shown in the middle of the figure below. In this section, we will install the package called plotly.

• •	RStudio				
💽 📲 🕼 🔚 🔚 🛛 📥 🛛 📣 Co to file/function	🔛 - Addins -				Project: (Non
guidance.Rmd ×	-0	Environment H	istory Connections		-
ःः। 🗊 🕞 🐬 🔍 🦽 Knit • ः 🔹 📲	insert •   🖓 👶   📑 Run •   🤹 •   🕿	🔐 🔒 🖙 Impo	ort Dataset + 🛛 🍕		🗏 List -   🔅
114		Global Environ	ment •		Q,
115 - ## RとRStudioを講義で使うための準備をす 116 117 - ### Rのパッケージを 118	5		Environm	ent is empty	
114:1 □ RとRStudioが動くことを確認する :	D Markelenen A				
Console Terminal × R Markdown × Jobs ×	Install Packages				
~/Google ドライブ/設業/バイオスタティスティクス/2020/gu	Install from: (7) Con	figuring Repositories			
	Repository (CRAN)	•			
l version 3.6.3 (2020-02-29) "Holding the W	Packages (separate multiple with space or	comma):	kages Help View	ver	-
opyright (C) 2020 The K Foundation for Statis Platform: x86 64-apple-darwin15.6.0 (64-bit)	papiot2	commay.	Delete 📻 Rename	🎲 More 🔹	
			e ドライブ > 授業 > パ	イオスタティスミ	Fィクス > 2020 > guidance
は、自由なソフトウェアであり、「完全に無保証」です。	Install to Library:			Size	Modified
ー定の条件に従えば、自由にこれを再配布することができます Fra条件の詳細に聞しては、'license(')' あるいは 'lice	/Library/Frameworks/R.framework/Versions/3.6/	lesources/library [C 2			
control and the treated where the	☑ Install dependencies		ntml	7.2 MB	Apr 9, 2020, 2:42 PM
は多くの貢献者による共同プロジェクトです。			Rmd	4.9 KB	Apr 9, 2020, 3:31 PM
洋しくは 'contributors()' と入力してください。			.Rmd	3.3 KB	Apr 9, 2020, 5:28 PM
citation()'と入力してください。	Insta	Cancel	e.html	7.2 MB	Apr 9, 2020, 2:59 PM
		🗌 🔍 guidance	e_e.Rmd	3.4 KB	Apr 9, 2020, 2:59 PM
demo() と入力すれはテモをみることができます。 'help() とすればオンラインヘルブが出ます。 help.stort() で HTML プラウザによるヘルプがみられ3 'q() と入力すれば R を終了します。	17.	🗌 📟 pics			
3 + 4 * 2					

3. The plotly package will be downloaded and installed.



4. Type the following commands to the "Console" to draw an interactive figure.

library(plotly)

plot\_ly(midwest, x = ~percollege, color = ~state, type = "box")

It will be shown in RStudio as follow.



## **Configuring the R Working Directory**

In R, it is often necessary to specify a "working directory" to run R. This enables R find R codes and data to use. Here, let's use RStudio to specify a working directory.

First, download the R code (readData.R) and data (iris.csv) from Google Drive below.

https://drive.google.com/drive/u/2/folders/160TTfRf0Lz9qxm8vLdciiaYZmS0Zr9BL

Next, move these files somewhere in a folder (directory) that is easy to work with for you. Here, I assume that you create a folder called Rtemp on your desktop and saved the files in the folder.

You can designate this folder as a "Working Directory" in RStudio as follows:

1. Click on the tab called "Files" in the lower right-hand corner of RStudio.

a Oli Con a Line of Cale Mathematica Con Address a			Theology (Misson)		
			Project. (Work)		
• readOuts.R ×	Environment History Connections		-0		
Contraction of the second seco	🗄 🤐 📊   🖙 Import Dataset +   🍯		Ust • 107		
<pre>A prod &lt;&gt; room.tsv( tris.tsv , rom.nomes = 1) A prod(dota) A preparation for the next week A required.packages &lt;&lt; c("KernSmoth", "mops", "mopdata", "plot 6 nem.packages &lt;&lt; required.packages[[(required.packages NinK in 7.m.(f()anoth/nam.norknone)).f 11 (f()plotee) = R.Sorge.</pre>	Environment is en	ngity	Q.		
Console Terminal × Jobs ×					
-10 d					
Copyright (C) 2020 The R Foundation for Statistical Computing Platform: x86_64-apple-dominid5.6.0 (64-bit) R は、自由なソフトウェアであり、「完全に最低圧」です。 一定の条件に気は、自由にこれを発展をすることができます。 配名条件の詳細に関しては、'license()' あるいは 'licence()' と入力してくださ い。 R は多くの異影者による共同プロジェクトです。 詳しくは 'contributors()' と入力してください。 また、R や R のパックージを出版物で引用する原の形式については 'citation()' と入力してください。 'demo()' と入力すればデモをみることができます。 'help() とすればデンテベルングがます。 'help() さればポンテインルングがます。 'help() さればポンテインルングがます。 'help() さん力すれば R を捨了します。 >	Kone Colder O Delan     More -     More - Deskapp     SRICYCLLEIN     SRICYCLLEIN     SRICYCLLEIN     SRICYCLLEIN     SRICYCLLEIN     SRICYCLLEIN     SI Archive Ini     Thumbuldb     SI     SI	5/2e 278 8 282 8 780.5 X8	Modfled Apr 9, 2020, 10:36 PM Apr 9, 2020, 12:34 AM Apr 9, 2020, 10:38 PM		

- 2. You can navigate to the various folders on your computer by clicking on a file path (here it is shown as "Home > Desktop") or folders shown in the "Files" window.
- 3. Here we want to move to "Rtemp", so click on "Rtemp". You can see that there is the files of R-code and data that you have saved there.
- 4. Once you have moved to the location where the R-code or data exist, click on "More" (next to the gear-shaped icon).



5. A drop down menu will then appear and select 'Set As Working Directory'.

l 🤍 🛡 an	RStudio	
🔹 🥸 🚰 • 🗧 🔂 👘 👘 Go to file/function 👘 🔛 • Addins •		🙁 Project: (None
Console Terminal × R Markdown × Jobs ×	Environment History Connections	
~/Desktop/ 🔅	🧹 🔄 📑 Import Dataset 🔹 🎻	🗏 List • 🛛 🕓
<pre>-/Desktop/ ⇒  version 3.6.3 (2020-02-29) "Holding the Windsock" opyright () 2020 The R Foundation for Statistical Computing latform: x86_64-apple-darwinI5.6.9 (64-bit)  it () fatavy&gt;トウェアであり、「完全に無保証」です。 = 二の条件に定えば、自由にこれを再発作することができます。 DE条件の評量に関しては、'license()' あるいは 'licence()' と入力してくだ  it は多くの貢献者による共同プロジェクトです。 #Lくは 'contributors()' と入力してください。 it よくの貢献者による共同プロジェクトです。 #Lくは 'contributors()' と入力してください。 it た、R や R のバッケージを出版者で引用する限の形式については citation()' と入力してください。 demo()' と入力すればデモをみることができます。 help()' と打しばオンラインヘルプが出ます。 help()' と入力すればデモをみることができます。 help()' と入力すればアモをみることができます。 help()' と入力すればアモをみることができます。 help()' と入力すれば R を終了します。 artite.csv(iris, "iris.csv", quote= F) setud("~/Desktop") setud("~/Desktop") </pre>		E List + ()

6. Now you designate a "Working Directory" to the "Rtemp" folder on your desktop.

This makes it easy to load data files to R.

First, let's open the R code. If you click on the file "readData.R" in the "Files" window, you will see the code in the top left window.



Let's select the first two lines of this and run them. To run, click on "Run". Alternatively, you can use "command + return (enter)" on Mac OS and "cntl + enter (return)" on Windows.

🔹 🧠 🖉 •   📄 🔛   📥   🏕 Co to file/function 🔤   🖾 • Addins							寒 Pro	ject: (None	
• readData.R ×		Environment	History	Connections				-	
Image: Source on Save     Q     P + 1     → Run   +   → Source +   ≥       1     data <- read.csv("tris.csv", row.notes = 1)		🖙 🔒 🖙 Import Dataset 👻 🎻					🗏 List • 🛛 💿		
		🐴 Global En	vironment •				Q.		
3:1 (Top Level) \$	R Script 0								
Console Terminal × Jobs ×		Files Plot	Package	s Help Vie	wer			- 6	
~/Desktop/Rtemp/ P		New Folder	er 🤨 Dele	te 🍺 Rename	🚭 More 🔹				
Copyright (C) 2020 The R Foundation for Statistical Computin	a	Home	> Desktop	> Rtemp		Size	Modified		
latform: x86_64-apple-darwin15.6.0 (64-bit)		£	anne.			212.6	NOUTLED		
2 は、自由なソフトウェアであり、「完全に無保証」です。		0 🖄 .Rh	story			852 B	Apr 9, 2020, 1	0:54 PM	
- 定の条件に従えば、自由にこれを再配布することができます。		iris.csv			4.1 KB	Apr 9, 2020, 1	0:14 PM		
88条件の詳細に関しては、'license()' あるいは 'licence()' と入力し	てください。	e readData.R			55 B	Apr 9, 2020, 1	0:52 PM		
: は多くの貢献者による共同プロジェクトです。 ましくは "contributors()" と入力してください。 また、R や R のバッケージを出感物で引用する際の形式については citation()" と入力してください。									
demo()、と入力すればデモをみることができます。 help()、とすればオンラインヘルブが出ます。 help.start()、で HTML プラグビよるヘルプがみられます。 (q()、と入力すれば 呆 を続了します。									
saturd/"_ (Darkton/Rhom")									

If it runs successfully, you'll see the following result.

	RStudio		
• 👒 🖆 • 🕞 🔂 📫 📝 Addins •			Project: (None)
readData.R ×	Environment History Connections		-0
🗢 🖉 📄 🔄 Source on Save 🔍 🥕 - 📄 👄 Run 📑 📑 Source - 🔍	🖙 🔒 🖙 Import Dataset 🔹 🍕		🗏 List + 🔯
<pre>data &lt;- read.csv("iris.csv", row.names = 1)</pre>	diobal Environment -		Q,
2 head(data) 3	Data		
	• data 150 obs. of 5 variable	5	
1 (Too Level) 2 R Script			
nsole Terminal × Jobs ×	Files Plots Packages Help Viewer		-
/Desktop/Rtemp/	💁 New Folder 🔍 Delete 📻 Rename 👙 More •		
は多くの貢献者による共同フロジェクトです。	A Home > Desktop > Rtemp		
→くは 'contributors()' と入力してください。	A Name	Size	Modified
itation()'と入力してください。	1		
	C     Shistory	852 B	Apr 9, 2020, 10:54 PM
mo() と入力すれはテモをみることかできます。 ele() とすればオンラインヘルプが出ます。	iris.csv	4.1 KB	Apr 9, 2020, 10:14 PM
elp.start()'で HTML ブラウザによるヘルプがみられます。 ()'と入力すれば R を終了します。	e readData.R	55 B	Apr 9, 2020, 10:52 PM
ashud/" (Daalisan (Daama")			
data <- read.csv("inis.csv", row.names = 1)	1		
read(data)			
sepal.length Sepal.midth Petal.length Petal.Width Species			
4.9 3.0 1.4 0.2 setosa			
4.7 3.2 1.3 0.2 setosa			
4.6 3.1 1.5 0.2 setosa			
5 6 7 5 1 6 7 Coboco			
5.4 3.9 1.7 0.4 setosa			

If you can do this far, there will be few issues in the class.