

Tomohisa OKADA

2023-07-05

Contents

1	Ready?	5
1.1	R(Studio)	5
1.2	Tips!	5
1.3	R(Studio)	5
2	Go!!	7
2.1	RStudio	7
2.2	7
2.3	R	8
2.4	9
2.5	10
3		11
3.1	11
3.2	12
3.3	14
3.4	15
3.5	16
4		17
4.1	17
4.2	22
4.3	23
4.4	24
4.5	24
5		25
5.1	25
5.2	25
5.3	25
6		27
6.1	27
6.2	28

6.3	29
7		31
7.1	31
7.2	31
7.3	32
7.4	33
7.5	34
7.6	34
7.7	35
7.8	36
7.9	37
7.10	38
8	{r} # library(wooldridge) # data("saving") # head(saving) # saving_modified <- # saving %>% # mutate(saving_rate = sav / inc) %>% # select(-black) %>% # arrange(-inc) # saving_modified #	39

Chapter 1

Ready?

1.1 R(Studio)

- R
- RStudio R R
- OK

1.2 Tips!

- R
1. R by
R
 2. R by &
Tidyverse R
 3. R by
 4. R by
R
 5. R by
tidyverse

1.3 R(Studio)

- (<https://posit.co/download/rstudio-desktop/>)
- 1:Install R R

- 2:Install RStudio RStudio
- R

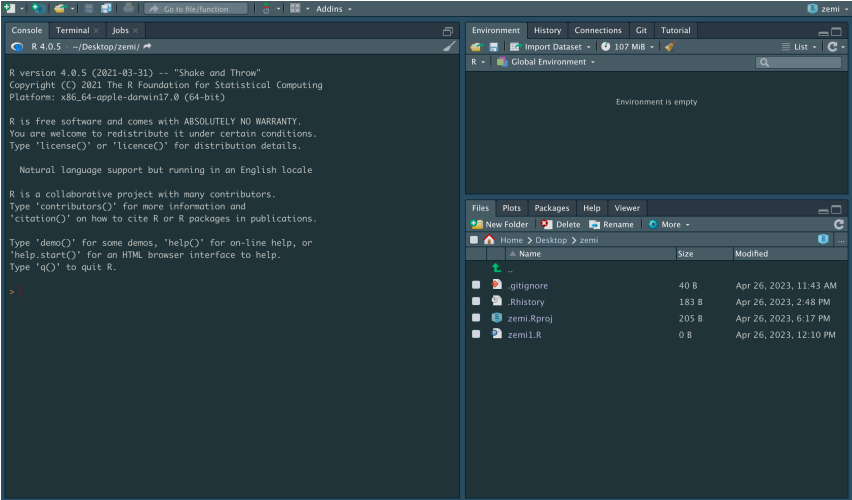
Chapter 2

Go!!

-

2.1 RStudio

- RStudio
-



2.2

-

- Console
 -
 - `> 1+1` **Enter mac return**
 - `[1] 2`
 - `2 1+1` `[1] 1`
-

2.3 R

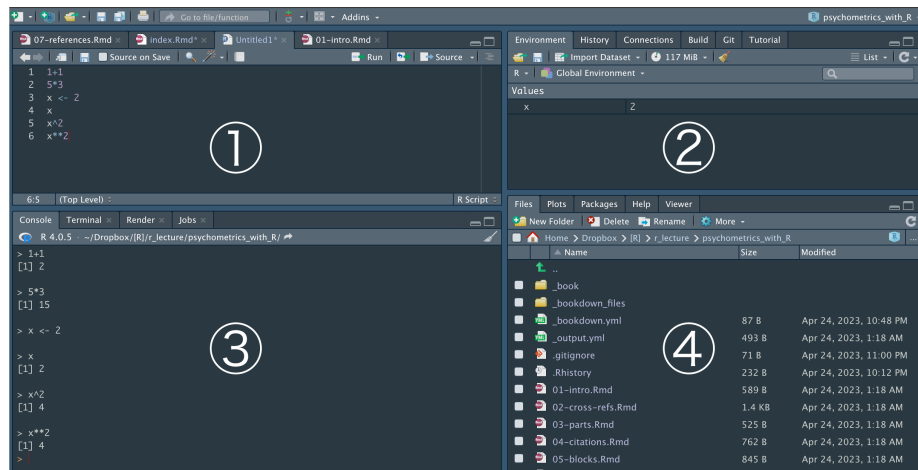
- Rstudio
- R

R

- RStudio R Script
- R
untitled1 R

RStudio

-



R

R

- untitled1
- 1 1+1 **ctrl+Enter** mac command+return
- ([1] 2)

- R **ctrl+s** mac command+s

- test.R
- R test.R ×
- R test.R
- R

-
- 2 5-2
- 2 **ctrl+Enter**
- 2 ([1] 3)
- 1 **ctrl+Enter**
- 1 ([1] 2)
- **ctrl+Enter**
- **ctrl+shift+Enter** mac command+shift+return
- **ctrl+Enter**

R

- 1.
- 2.
- 3.
- ...

2.4

- R
- R
- New Directry → New Project

- Create Project
-
- ()
- .Rproj
- mac Document sugoi_project
-
-
- 1. .Rproj Rstudio
- 2.

2.5

- zemi
 - zemi
 - zemi.Rproj
- ※ zemi

Chapter 3

- 1. in \mathbb{R}
 2. in \mathbb{R}
- \mathbb{R}

3.1

`:` +

```
1 + 1
```

```
## [1] 2
```

`:` -

```
5 - 2
```

```
## [1] 3
```

`:` *

```
4 * 5
```

```
## [1] 20
```

`:` /

```
8 / 2
```

```
## [1] 4
```

```

: ^** 42
4 ^ 2

## [1] 16
4 ** 2

## [1] 16
• 9÷2 4 1
• 1
: %/%

```

```
9%/%2
```

```
## [1] 4
mod : %%

```

```
9%%2
```

```
## [1] 1
```

3.2

```

• 1
•
•
•
x <- 1 # <-
# #
#
#
• x 1
x OK
• OK
•
x

```

```
## [1] 1
```

```
•
y <- 1
z <- 2
y * z # 1*2
```

```
## [1] 2
```

```
•
x <- 1 #x 1
x
```

```
## [1] 1
```

```
x <- 2 #x 2
x
```

```
## [1] 2
```

```
•
x <- 2+5
x # 7
```

```
## [1] 7
```

```
•
•
suuji <- 2 #suuji 2
suuji <- suuji + 1 #suuji=2 1      suuji
suuji # suuji
```

```
## [1] 3
```

```
•
•   moji
#   "   "
#   '   '
#
moji <- "   "
moji
```

```
## [1] "   "
```

```
• x 3 y 6
```

```
• x y 2 45
```

3.3

- R
- `sqrt()`
- `()`

```
sqrt(2)
```

```
## [1] 1.414214
```

- `xxx()`
- `()`
-
- `sqrt(2) 2 1.414214`

-
- `log()`
- `10`

```
log(10)
```

```
## [1] 2.302585
```

- `2 base=10 10`

```
log(10, base = 10)
```

```
## [1] 1
```

- `(numeric)`

```
( x )
```

```
log(x)
```

```
log(x, base=y)
```

```
y
```

```
sqrt(x)
```

```
x
```

```
exp(x)
```

```
x  $e^x$ )
```

```
abs(x)
```

```
x
```

```
round(x,y)
```

```
y x IEEE754 *
```

```
floor(x)
```

```
x
```

```
ceiling(x)
```

```
x
```

-

- `help()`
 - `() help()`
 - `log() help(log)`
 - Rstudio
-

3.4

- `+ 1`

```
suuji <- 2
suuji + 1
```

```
moji <- " "
moji + 1
```

- `3`
- `suuji 2 numeric double`
- `moji character`
- `character +`
- `+`
- `typeof() mode()`

```
typeof(suuji)
```

```
## [1] "double"
```

```
typeof(moji)
```

```
## [1] "character"
```

- `TRUE FALSE logical`
-

- `suuji moji mode(suuji) mode(moji)`
 - `typeof(suuji) mode(suuji)`
-

3.5

- zemi exercise_ch3.R R
- exercise_ch3.R

1. `abs(-5)`

2. `x <- exp(10) log(x)`

1. `round(0.45, 1)` 0.5 0.4
 IEEE 5

2. `floor(0.45 * 10)/10 0.5`
 x y `floor(x * 10y + 0.5)/10y`

Chapter 4

- 1. in \mathbb{R}
 2. in \mathbb{R}
- \mathbb{R}

4.1

- \mathbb{R}

4.1.1

- 1
- `c()`
-
- 5 2,4,2,3,5 v

```
v <- c(2, 4, 2, 3, 5) #  
v # v
```

```
## [1] 2 4 2 3 5
```

- 2,3,4,5,6

```
v2_6 <- c(2:6) # n:m n m  
v2_6
```

```
## [1] 2 3 4 5 6
```

```

•      =
•
•
v

## [1] 2 4 2 3 5
v+2 #

## [1] 4 6 4 5 7
2*v #

## [1] 4 8 4 6 10

```

```

• v v-2 v/2 v^2

```

```

• R
• 2
v1 <- c(1, 2)
v2 <- c(2, 4)

```

```

• +
• 1 2 (2 4)
v1 + v2

```

```
## [1] 3 6
```

```

• * 1 2
v1 * v2

```

```
## [1] 2 8
```

```

•
• R %*%
v1 %*% v2

```

```
##      [,1]
## [1,]    10
```

```

• =
# v1 3 v2 2
# v2 1 v1

```

```
v1 <- c(1, 2, 3) #3
v2 <- c(2, 4) #2
v1 + v2 #
```

```
## Warning in v1 + v2: longer object length is not a multiple of shorter object
## length
```

```
## [1] 3 6 5
```

- `length()`

```
length(v1)
```

```
## [1] 3
```

- `v1 v2 - / ^`

- `1 2`

- `n`

- `x n x[n]`

- `v=c(2,4,2,3,5)`

- `2 4 v[2]`

- `2 (4,2,3) v[2:4]`

```
v[2:4]
```

```
## [1] 4 2 3
```

4.1.2

- 2
- $2 \ 3 \ \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \ M \ \text{matrix()}$

```
# 1:6 c(1:6) 1,2,3,4,5,6
```

```
# 6 2 (row) 3 (col)
```

```
#byrow = T 6 z
```

```
#byrow = T
```

```
M <- matrix(1:6, nrow = 2, ncol = 3, byrow = T)
```

```
M
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
```

•

```
v1 <- c(1,2,3)
v2 <- c(1,1,1) #2 v1 v2
```

```
rbind(v1, v2) #v1 v2 (row)
```

```
##      [,1] [,2] [,3]
## v1     1    2    3
## v2     1    1    1
```

```
cbind(v1, v2) #v1 v2 (column)
```

```
##      v1 v2
## [1,]  1  1
## [2,]  2  1
## [3,]  3  1
```

```
rbind(M, v1) #
```

```
##      [,1] [,2] [,3]
##      1    2    3
##      4    5    6
## v1     1    2    3
```

• x x[]
• M

```
# 2 1 M21
M21 <- M[2,1]
M21
```

```
## [1] 4
```

```
# 2
M[2,]
```

```
## [1] 4 5 6
```

```
# 1
M[,1]
```

```
## [1] 1 4
```

```
# 1,2 1,3
M[c(1,2),c(1,3)]
```

```
##      [,1] [,2]
## [1,]    1    3
## [2,]    4    6
```

```
• n n      + -
•      %*%
```

```
M # 2 3
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
```

```
M2 <- matrix(c(1,2,0,1,0,2), nrow = 3, ncol = 2, byrow = T) # 3 2
M2 # 2 3
```

```
##      [,1] [,2]
## [1,]    1    2
## [2,]    0    1
## [3,]    0    2
```

```
# MN
M %*% M2
```

```
##      [,1] [,2]
## [1,]    1   10
## [2,]    4   25
```

```
# MN
M2 %*% M
```

```
##      [,1] [,2] [,3]
## [1,]    9   12   15
## [2,]    4    5    6
## [3,]    8   10   12
```

```
v <- c(1,2,3) #
```

```
# %*%
#
```

```
v %*% M2 # M2%*%v
```

```
##      [,1] [,2]
## [1,]    1   10
```

4.2

- `rbind()` `cbind()`
-

-

(x)	
<code>summary(x)</code>	
<code>max(x)</code>	x
<code>min(x)</code>	x
<code>mean(x)</code>	x
<code>median(x)</code>	x
<code>var(x)</code>	x
<code>sd(x)</code>	x
<code>sum(x)</code>	x
<code>range(x)</code>	x
<code>length(x)</code>	x
<code>sort(x)</code>	x
<code>sort(x, decreasing = TRUE)</code>	x

- 5

```
age <- c(36, 16, 43, 18, 22) #5
```

- age

```
summary(age) #
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##       16      18      22      27      36      43
```

```
max(age) #      5      43
```

```
## [1] 43
```

```
mean(age) #      5      27
```

```
## [1] 27
```

```
var(age) #      5      141
```

```
## [1] 141
```

- age `min()`, `median()`, `sd()`, `sum()`, `range()`, `length()`, `sort()`

-
- M
-
-

(x)		
matrix(0, nrow=2, ncol=3)	2 3	
diag(5)	5×5	
diag(X) <- 1	X	1
t(X)	X	
solve(X)	X	
det(X)	X	
rowSums(X)	X	
colSums(X)	X	
RowMeans(X)	X	
colMeand(X)	X	

- M

4.3

3.7 OK

- R
-
-
-

```
<- function ( ) {  
  
}
```

- $a/(1-x)$ a x
- inf_geo()

```
inf_geo <- function (a, x) {  
  a/(1-x)  
}
```

```

•
#a=1 x=0.8
inf_geo(1, 0.8)

## [1] 5

```

4.4

- R
 - R
 -
 -
 -
-

4.5

- zemi exercise_ch4.R R
- exercise_ch4.R
- a 149cm, b 153cm, c 169cm, d 174cm
- a 36kg, b 48kg, c 61kg, d 65kg

1. 4 h

2. h

3.
1cm 0.39

4. 4 w
5. h w 2×4 M

6. M b

7. w h2

1. x y siswa_gonyu(x,y)

2. siswa_gonyu(0.4445,3)

Chapter 5

- R

5.1

-
-
- R

5.2

-
- `install.packages()`
- tidyverse¹
- tidyverse R

```
#  
install.packages("tidyverse") # " "
```

5.3

-
- tidyverse

```
# RStudio  
library(tidyverse) # " "
```

- `install.packages()`

¹

- `library()` Rstudio
- R `library()`

Chapter 6

-
- 1. in R
- R

6.1

-
- 4

Name	Age	Height	Weight	Gender
Tanaka	10	149.5	36	male
Suzuki	18	153	65	female
Okada	41	171	58	male
Watanabe	26	174.5	127	male

- R 1

6.1.1

- data.frame()

```
#
name <- c("Tanaka", "Suzuki", "Okada", "Watanabe")
age <- c(10, 18, 36, 23) #
height <- c(149.5, 153.0, 171.0, 174.5)
weight <- c(36, 65, 58, 127)
```

1

```
gender <- c("male", "female", "male", "male")

# data.frame()
#           df
df <- data.frame(name, age, height, weight, gender)
df
```

```
##      name age height weight gender
## 1 Tanaka  10  149.5     36   male
## 2 Suzuki  18  153.0     65 female
## 3 Okada   36  171.0     58   male
## 4 Watanabe 23  174.5    127   male
```

6.1.2

```
• df      age  $
df$age
```

```
## [1] 10 18 36 23
```

```
• df$age
toshi <- df$age
toshi
```

```
## [1] 10 18 36 23
```

6.2

```
•
•
• CSV .csv Excel .xlsx, .xls
•
```

6.2.1

```
•
• R
• R
• 2
```

```
•
•
```

6.2.2 csv

- csv read.csv() ³
- sokutei.csv csv data
- head()

```
data <- read.csv("sokutei.csv")
```

```
head(data) #head()
```

6.2.3 Excel

- Excel readxl
- readxl read_excel()
- sokutei.xls sokutei

```
install.packages("readxl") #
library(readxl)
sokutei <- read_excel("sokutei.xls")
```

6.2.4

-
-
- data
- sokutei.csv read.csv("data/sokutei.csv")

6.3

1. zemi data
2. sokutei.csv data
3. sokutei_csv
4. head() sokutei_csv
5. sokutei.xls data
6. sokutei_excel
7. head() sokutei_excel
8. sokutei_excel weight w

³R ver4.1.0 tidyverse |> %>%

9. `w`

- `Wooldridge`
- `wooldridge`

-

1. `wooldridge`

2. `wooldridge`

3. `data("wage1")` `wooldridge` `wage1` ⁴

4. `head(wage1)` `wage1`

5. `help(package="wooldridge")` `wage1`

⁴ `TRUE` `FALSE` `filter(bmi >= 19)` `TRUE`

Chapter 7

-
- 1. tidyverse in R

7.1

-
-
- tidyverse
- tidyverse ¹

```
library(tidyverse)
```

7.2

- tidyverse (%>) ^{2 3}
- R
- x

```
x <- c(1,2,3,4,5)
```

```
sd(x)
```

```
## [1] 1.581139
```

-

¹

² tidyverse magrittr

³R ver4.1.0 tidyverse |> %>%

```
x %>% sd()
```

```
## [1] 1.581139
```

```
•
•
•      df
```

```
df %>% head() #
```

```
##      name age height weight gender
## 1  Tanaka  10  149.5     36   male
## 2  Suzuki  18  153.0     65 female
## 3   Okada  36  171.0     58   male
## 4 Watanabe 23  174.5    127   male
```

```
•
```

```
read.csv("data/sokutei.csv") %>% head()
```

```
##      Name Age Height Weight Gender
## 1  Tanaka  10  149.5     36   male
## 2  Suzuki  18  153.0     48 female
## 3   Okada  41  171.0     58   male
## 4 Watanabe 26  174.5     65   male
## 5     Sato 32  159.0     54 female
## 6 Takahashi 16  169.0     87   male
```

```
•
```

```
•
```

```
sd_x <-
  x %>%
  var() %>% #x
  sqrt()   #
#
# sd(x)
sd_x
```

```
## [1] 1.581139
```

```
•
```

```
•
```

7.3

```
• tidyverse dplyr
•
```


- dplyr select()
- df name

```
df %>%
  select(age, height, weight, gender)
```

```
##   age height weight gender
## 1  10  149.5    36   male
## 2  18  153.0    65 female
## 3  36  171.0    58   male
## 4  23  174.5   127   male
```

- df name
- select() - !

```
df %>%
  select(-name) # - !
```

```
##   age height weight gender
## 1  10  149.5    36   male
## 2  18  153.0    65 female
## 3  36  171.0    58   male
## 4  23  174.5   127   male
```

7.4

- rename()
- df age toshi

```
df %>%
  rename(toshi = age) # =
```

```
##   name toshi height weight gender
## 1 Tanaka    10  149.5    36   male
## 2 Suzuki    18  153.0    65 female
## 3 Okada     36  171.0    58   male
## 4 Watanabe  23  174.5   127   male
```

- select()
- df name age toshi

```
df %>%
  select(toshi = age, height, weight, gender)
```

```
##   toshi height weight gender
## 1    10  149.5    36   male
## 2    18  153.0    65 female
## 3    36  171.0    58   male
## 4    23  174.5   127   male
```

7.5

-
- `dplyr` `filter()`
- `age 18`

```
df %>%
  filter(age >= 18) # filter()
```

```
##      name age height weight gender
## 1  Suzuki  18  153.0     65 female
## 2   Okada  36  171.0     58  male
## 3 Watanabe 23  174.5    127  male
```

7.6

- `age >= 18 bmi 19`
- ⁴

>	a > b	a b
>=	a >= b	a b
<	a < b	a b
<=	a <= b	a b
==	a == b	a b
!=	a != b	a b
%in%	a %in% c(a, b, c)	a c(a, b, c)

!	!(a==b)	a b
&	a & b	a b
&&	a && b	a b
	a b	a b
	a b	a b

- `df`
- `select()`

1. `age 18 weight 60kg`
- 2.
- 3.

⁴ TRUE FALSE `filter(bmi >= 19)` TRUE

7.7

- `dplyr` `mutate()`
- `df` BMI
BMI $\text{kg} \div (\text{m})^2$

```
df <- df %>%
  mutate(bmi = weight / (height/100)^2) #BMI df
df
```

```
##      name age height weight gender    bmi
## 1 Tanaka  10  149.5     36   male 16.10720
## 2 Suzuki  18  153.0     65 female 27.76710
## 3 Okada   36  171.0     58   male 19.83516
## 4 Watanabe 23  174.5    127   male 41.70738
```

2

- `age` 18 1 18 0 `is_child`
- `mutate()` `if_else()`

```
df <- df %>%
  mutate(is_child =
    if_else(
      age < 18, #if_else()
      1, #2 TRUE
      0 # FALSE
    ))
df
```

```
##      name age height weight gender    bmi is_child
## 1 Tanaka  10  149.5     36   male 16.10720      1
## 2 Suzuki  18  153.0     65 female 27.76710      0
## 3 Okada   36  171.0     58   male 19.83516      0
## 4 Watanabe 23  174.5    127   male 41.70738      0
```

- `if_else()` 3
- 3 `case_when`
- (WHO) BMI

 BMI

BMI		
16.00	16.99	
17.00	18.49	
18.50	24.99	
25.00	29.99	
30.00	34.99	(1)
35.00	39.99	(2)
40.00		(3)

- fat

```
df <- df %>%
  mutate( fat =
    case_when(
      bmi < 16 ~ " ",
      bmi >= 16 & bmi < 17 ~ " ",
      bmi >= 17 & bmi < 18.5 ~ " ",
      bmi >= 18.50 & bmi < 25 ~ " ",
      bmi >= 25 & bmi < 30 ~ " ",
      bmi >= 30 & bmi < 35 ~ " (1)",
      bmi >= 35 & bmi < 40 ~ " (2)",
      # TRUE~
      TRUE~" (3)" # bmi >= 40 ~ " (3)",
    ))

df
```

```
##      name age height weight gender      bmi is_child      fat
## 1  Tanaka  10  149.5     36   male 16.10720         1
## 2  Suzuki  18  153.0     65 female 27.76710         0
## 3   Okada  36  171.0     58   male 19.83516         0
## 4 Watanabe 23  174.5    127   male 41.70738         0 (3)
```

7.8

- dplyr arrange()
- df age

```
df <- df %>% arrange(age)

#      arrange(-age)

df
```

```
##      name age height weight gender      bmi is_child      fat
## 1  Tanaka  10  149.5     36   male 16.10720         1
```

```
## 2 Suzuki 18 153.0 65 female 27.76710 0
## 3 Watanabe 23 174.5 127 male 41.70738 0 (3)
## 4 Okada 36 171.0 58 male 19.83516 0
```

7.9

- dplyr summarise() ⁵
-
- df age mean_age

```
df %>%
  summarise(mean_age = mean(age))
```

```
## mean_age
## 1 21.75

• mean(df$age)
• summarise() group_by()
• group_by()
•
```

```
df %>%
  group_by(gender) %>% #
  summarise(mean_age = mean(age)) #
```

```
## # A tibble: 2 x 2
##   gender mean_age
##   <chr>     <dbl>
## 1 female      18
## 2 male       23
```

- summarise()
-

```
df %>%
  group_by(gender) %>% #
  summarise(n = n(),
            height_mean = mean(height),
            height_median = median(height),
            height_sd = sd(height),
            #
            # (NA)
            )
```

```
## # A tibble: 2 x 5
##   gender      n height_mean height_median height_sd
##   <chr> <int>     <dbl>         <dbl>     <dbl>
```

⁵summarize()

## 1 female	1	153	153	NA
## 2 male	3	165	171	13.5

7.10

```

1. zemi      data

2. wooldridge

3. data("saving")  wooldridge  saving
4. head(saving)    saving
5. saving saving_rate      sav/inc
6. saving black
7. saving inc
8.          saving_modified

```

Chapter 8

```
{r} # library(wooldridge)
# data("saving") #
head(saving) #
saving_modified <- #
saving %>% #
mutate(saving_rate = sav /
inc) %>% #
select(-black) %>% #
arrange(-inc) #
saving_modified #
```