R

2017-09-19

Contents

Chapter 1

“ R”! R,
. .

:

- R
- RStudio

, . , RStudio ,
R. .

1.1

( ), : **HTML** **R**.

1.1.1 HTML

HTML *1*, *2* . .
. , ,
.

:

```
a <- 5
sin(sqrt(a))
## [1] 0.7867491
sqrt(sin(a) + 2)
## [1] 1.020331
```

`sin(sqrt(a))` , *## [1] 0.7867491* — ,
(RStudio).

HTML ,
( ) , (
).

1.1.2 R

R *1*, *2* . **R** . . , **R** - . , , , , **HTML**. R , .

R , . , - , , .

Windows *CP1251* , OS X — *UTF8*. , ( ). — , RStudio. , *File > Open With Encoding...*

1.2

:

- : *Run* `Ctrl+Enter` (`Cmd+Enter` OS X).
- : *Run* `Ctrl+Enter` (`Cmd+Enter` OS X).
- :
  - *Run* `Ctrl+Enter` (`Cmd+Enter` OS X)
  - `Ctrl+Alt+Enter` (`Cmd+Alt+Enter` OS X)
  - *Source*

  Source `Ctrl+Alt+Enter` , , (Windows 10+ OS X, , Unicode).

, *Code > Run Region*

, , . , . , - , , . , — . . .

1.3

— , `#`. . , .

, . , - — .

`#` . , . , .

.
(-) (_), .
. . . ,
— . :

```
#
# ---------------------------
# , , , , , :

a <- 3 + 2 #
b <- 4 ^ 8 #
c <- b %% a #

#
d <- c / a

#
e <- d * b
```

, , , .
, , .

, , *Code > Insert Section*
`Ctrl+Shift+R` (`Cmd+Shift+R` OS X)

1.4 R

, R.
, .
. ( )
(*style guides*)
R, R Google,
.

. ,
( ) : , , , -
.

1.5

R, . ,
,

`$`
`#`
`&`
`/`
`\`
`|`
`^`
`@`

~
''
""
``

Chapter 2

R,

2.1

— R. *c* . R ,
. , — .
, :

```
2 + 3
## [1] 5
2 - 3
## [1] -1
2 * 3
## [1] 6
```

, :

```
2.5 + 3.1
## [1] 5.6
```

. (**) (^),
“ ”:

```
2 ^ 3
## [1] 8
2 ** 3
## [1] 8
```

:

```
5 / 3
## [1] 1.666667
5 / 2.5
## [1] 2
```

( ) %/%:

```
5 %/% 3
## [1] 1
```

%%:

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

, . “ ”.
— . , . .
:

```
5%/%3
## [1] 1
```

```
5 %/% 3
## [1] 1
```

, .
3. , , , 3 4.
. , , - . , ,
? . ,
. <-

```
a <- 5
b <- 3
```

, :

```
a
## [1] 5
b
## [1] 3
```

:

```
a + b
## [1] 8
a - b
## [1] 2
a / b
## [1] 1.666667
a %/% b
## [1] 1
a %% b
## [1] 2
```

3 4 .

```
b <- 4
a + b
## [1] 9
a - b
## [1] 1
a / b
## [1] 1.25
a %/% b
## [1] 1
a %% b
## [1] 1
```

, , !

:

```
c <- b
d <- a+c
```

, :

```
c
## [1] 4
d
## [1] 9
```

:

```
e <- d + 2.5
e
## [1] 11.5
```

- :

```
f <- -2
f
## [1] -2
f <- -e
f
## [1] -11.5
```

, , , .
2. , . c 4, d 9:

```
c %% 2
## [1] 0
d %% 2
## [1] 1
```

2.1.1

,
. , . , Z, 2 : x y.
Z(x,y). , x y . ,
, , . .

— . sqrt(x), abs(x), sin(x), cos(x),
tan(x) asin(y), acos(y), atan(y) . , ,
:

```
sqrt(a)
## [1] 2.236068
sin(a)
## [1] -0.9589243
tan(1.5)
## [1] 14.10142
abs(a + b - 2.5)
## [1] 6.5
```

, :

```
sin(sqrt(a))
## [1] 0.7867491
sqrt(sin(a) + 2)
## [1] 1.020331
```

, :

```
b <- sin(sqrt(a))
b
## [1] 0.7867491
```

b , . , . ,
, a — , , :

```
b <- sin(a)
b
## [1] -0.9589243
b <- abs(b)
b
## [1] 0.9589243
```

2.2

— . . , :

```
s <- "          ,          ( .     )"
s
## [1] "          ,          ( .     )"
```

nchar()

```
nchar(s)
## [1] 56
```

. .
. , +, paste(). ,

```
s1 <- "          ,"
s2 <- "          "
s3 <- "( .     )"
```

:

```
s1
## [1] "          ,"
s2
## [1] "          "
s3
## [1] "( .     )"
```

:

```
s <- paste(s1, s2)
s
## [1] "          ,          "
s <- paste(s1, s2, s3)
s
## [1] "          ,          ( .     )"
```

( ) ,
( ). , 1950
1850 . . year, pop. ,
. “ ”” ? .

,

```
year <- 1950
pop <- 1850
```

```
s1 <- "                                  "
s2 <- "          "
s3 <- "  .   "
s <- paste(s1, year, s2, pop, s3)
s
## [1] "                                  1950           1850   .   "
```

. . ,

.

2.3

. .

. . , . as.Date().

— , - . ,

. ( ):

```
birth <- as.Date('1986/02/18')
birth
## [1] "1986-02-18"
```

Sys.Date():

```
current <- Sys.Date()
current
## [1] "2017-09-19"
```

. , , , .. ,

:

```
livedays <- current - birth
livedays
## Time difference of 11536 days
```

. , , 40 :

```
current + 40
## [1] "2017-10-29"
```

.

2.4 ( )

, . TRUE ( ) FALSE ( ).

T F .

:

- (==) —
- (!=) —
- (<) —

- (<=) —
- (>) —
- (>=) —

, :

```
a <- 1
b <- 2
a == b
## [1] FALSE
a != b
## [1] TRUE
a > b
## [1] FALSE
a < b
## [1] TRUE
```

, . :

- (&&) -
- (||) -
- (!) - (,)

```
c<-3
(b>a) && (c>b)
## [1] TRUE
(a>b) && (c>b)
## [1] FALSE
(a>b) || (c>b)
## [1] TRUE
!(a>b)
## [1] TRUE
```

if

Chapter 3

—
.
R. , , ,

3.1

. , , ,
. , .

. :

1.
2. (“ ”),
3.

c():

```
#          -
colors <- c("    ", "    ", "   ", "      ", "     ")
colors
## [1] "    "    "    "    "   "      "     " "     "
```

```
#          -              (              )
lengths <- c(28, 40, 45, 19, 38)
lengths
## [1] 28 40 45 19 38
```

```
#                -                      (              )
opens <- c(FALSE, TRUE, TRUE, FALSE, FALSE)
opens
## [1] FALSE  TRUE  TRUE FALSE FALSE
```

, , , — ,
. .
, vector(). 2 :

- mode "logical", "integer", "numeric" ("double"), "complex", "character" "raw"
- length

:

```
#       5      ,
intvalues <- vector(mode = "integer", length = 5)
intvalues #
## [1] 0 0 0 0 0

#       10      ,                                  (    )
charvalues <- vector("character", 10)
charvalues #
##  [1] "" "" "" "" "" "" "" "" "" ""
```

, = , .
. . , ,
, .

, . ,
?vector

, — . M N,
: M:N:

```
index <- 1:5 #         c(1,2,3,4,5)
index
## [1] 1 2 3 4 5
index <- 2:4 #         c(2,3,4)
index
## [1] 2 3 4
```

— seq(), / :

```
seq(from = 1, by = 2, length.out = 10) # 10         ,
##  [1]  1  3  5  7  9 11 13 15 17 19
seq(from = 2, to = 20, by = 3) #   2   20      3 (          )
## [1]  2  5  8 11 14 17 20
seq(length.out = 10, to = 2, by = -2) #                  10                    ,     2
##  [1] 20 18 16 14 12 10  8  6  4  2
```

, seq() (,
(=). , . , ,
.

:

```
seq(from = as.Date('2016/09/01'), by = 1, length.out = 7) #                  2016/2017
## [1] "2016-09-01" "2016-09-02" "2016-09-03" "2016-09-04" "2016-09-05"
## [6] "2016-09-06" "2016-09-07"

seq(from = Sys.Date(), by = 7, length.out = 5) #              ,
## [1] "2017-09-19" "2017-09-26" "2017-10-03" "2017-10-10" "2017-10-17"
```

3.2

:

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

```
colors[3] #
## [1] "   "
```

( ) length():

```
length(colors)
## [1] 5
```

, :

```
n <- length(colors)
colors[n]
## [1] "     "
```

. , 4 . :

```
lengths[1:4]
## [1] 28 40 45 19
```

. , :

```
m <- 1
n <- 4
index <- m:n
lengths[index]
## [1] 28 40 45 19
```

. ,
, . :

```
index <- c(1, 3, 4) #          1, 3  4
lengths[index]
## [1] 28 45 19

index <- c(5, 1, 4, 2) #
lengths[index]
## [1] 38 28 19 40
```

3.3

. , — , ,
, , , , , :

```
min(lengths) #
## [1] 19
max(lengths) #
## [1] 45
range(lengths) #          =      -
## [1] 19 45
mean(lengths) #
## [1] 34
median(lengths) #
## [1] 38
var(lengths) #       (        -     , variation)
## [1] 108.5
sd(lengths) #                   (standard deviation)
```

```
## [1] 10.41633
sum(lengths) #
## [1] 170
```

R . R ,
1 (11). , , , :

```
lengths * 1000 #
## [1] 28000 40000 45000 19000 38000
sqrt(lengths) #
## [1] 5.291503 6.324555 6.708204 4.358899 6.164414

stations <- c(20, 21, 22, 12, 24) #

dens <- stations / lengths #                    =   -      /
dens
## [1] 0.7142857 0.5250000 0.4888889 0.6315789 0.6315789
```

3.4

:

```
lengths2 <- sort(lengths) #
lengths2 #
## [1] 19 28 38 40 45
lengths #
## [1] 28 40 45 19 38

lengths2 <- sort(lengths, decreasing = TRUE) #                    .              decreasing
lengths2 #
## [1] 45 40 38 28 19
lengths #
## [1] 28 40 45 19 38
```

— . , , ( )
. , , max(lengths). ,
, . , ,
lengths , .
match():

```
l <- max(lengths) #
idx <- match(l, lengths) #              ,      l,       lengths
color <- colors[idx] #
color
## [1] "   "
```

, !

```
s <- paste(color, "              -        .                ", l, " ")
s
## [1] "                  -        .                45  "
```

“ ”” __ :

```
colors[match(max(dens),dens)]
## [1] "    "
```

Chapter 4

,

, .

4.1

— 2 . ,
. , , (, , . .).
, . , .

, , matrix, 3 : ,
:

```
v <- 1:12  #                          1   12
m <- matrix(v, nrow = 3, ncol = 4)
m
##      [,1] [,2] [,3] [,4]
## [1,]    1    4    7   10
## [2,]    2    5    8   11
## [3,]    3    6    9   12
```

, . ,
byrow = TRUE:

```
m <- matrix(v, nrow = 3, ncol = 4, byrow = TRUE)
m
##      [,1] [,2] [,3] [,4]
## [1,]    1    2    3    4
## [2,]    5    6    7    8
## [3,]    9   10   11   12
```

, :

```
m[2,4]  # 2    , 4
## [1] 8
m[3,1]  # 3    , 1
## [1] 9
```

, . ,
. :

```
m[2,]  # 2
## [1] 5 6 7 8
m[,3]  # 3 c
## [1]  3  7 11
```

, :

```
log(m)  #
##          [,1]      [,2]     [,3]     [,4]
## [1,] 0.000000 0.6931472 1.098612 1.386294
## [2,] 1.609438 1.7917595 1.945910 2.079442
## [3,] 2.197225 2.3025851 2.397895 2.484907
sum(m)  #
## [1] 78
median(m) #
## [1] 6.5
```

:

```
sort(m)
##  [1]  1  2  3  4  5  6  7  8  9 10 11 12
```

, , :

```
t(m)  #
##      [,1] [,2] [,3]
## [1,]    1    5    9
## [2,]    2    6   10
## [3,]    3    7   11
## [4,]    4    8   12
m2<-matrix(-3:3,nrow = 3, ncol = 3)
## Warning in matrix(-3:3, nrow = 3, ncol = 3):         [7]
##                  [3]
m2
##      [,1] [,2] [,3]
## [1,]   -3    0    3
## [2,]   -2    1   -3
## [3,]   -1    2   -2
det(m2) #
## [1] -21
det(m)  #    !
## Error in determinant.matrix(x, logarithm = TRUE, ...): 'x' must be a square matrix
```

%*%. , ,
:

```
m2 %*% m
##      [,1] [,2] [,3] [,4]
## [1,]   24   24   24   24
## [2,]  -24  -28  -32  -36
## [3,]   -9  -10  -11  -12
m %*% m2  #    !
## Error in m %*% m2:
```

match(), , . which().
m 8, :

```
which(m == 8, arr.ind = TRUE)
##      row col
## [1,]   2   4
```

, 1×2. , .
, :

```
indexes <- which(m == 8, arr.ind = TRUE)
row <- indexes[1,1]
col <- indexes[1,2]
m[row,col]
## [1] 8
```

! 8.

— , . cbind() rbind(). . :

```
lengths <- c(28, 40, 45, 19, 38)
stations <- c(20, 21, 22, 12, 24)
cbind(lengths, stations)  #
##      lengths stations
## [1,]      28       20
## [2,]      40       21
## [3,]      45       22
## [4,]      19       12
## [5,]      38       24
rbind(lengths, stations)  #
##          [,1] [,2] [,3] [,4] [,5]
## lengths    28   40   45   19   38
## stations   20   21   22   12   24
```

C :

```
mm <- cbind(lengths, stations)
mm[,2]/mm[,1]  #              1
## [1] 0.7142857 0.5250000 0.4888889 0.6315789 0.6315789
```

:

```
dens <- mm[,2]/mm[,1]
mm<-cbind(mm, dens)
mm
##      lengths stations      dens
## [1,]      28       20 0.7142857
## [2,]      40       21 0.5250000
## [3,]      45       22 0.4888889
## [4,]      19       12 0.6315789
## [5,]      38       24 0.6315789
```

Environment

Data

, , , , - , ( ).
, , ,
:

```
colors <- c("    ", "    ", "   ", "      ", "     ")
mm2<-cbind(mm,colors)
```

```
mm2  #              ,
##      lengths stations dens                colors
## [1,] "28"    "20"     "0.714285714285714" "    "
## [2,] "40"    "21"     "0.525"             "    "
## [3,] "45"    "22"     "0.488888888888889" "   "
## [4,] "19"    "12"     "0.631578947368421" "      "
## [5,] "38"    "24"     "0.631578947368421" "     "
```

, :

```
mm2[,2]/mm2[,1]
## Error in mm2[, 2]/mm2[, 1]:
```

4.2

— . . R
(*data frame*). data.frame():

```
t<-data.frame(colors,lengths,stations)
t  #        ,
##       colors lengths stations
## 1              28       20
## 2              40       21
## 3               45       22
## 4             19       12
## 5              38       24
```

:

```
t<-cbind(t, dens)
t
##       colors lengths stations      dens
## 1              28       20 0.7142857
## 2              40       21 0.5250000
## 3               45       22 0.4888889
## 4             19       12 0.6315789
## 5              38       24 0.6315789
```

data.frame() cbind() . ,
, — :

, :

```
t[2,2]
## [1] 40
t[,3]
## [1] 20 21 22 12 24
t[4,]
##       colors lengths stations      dens
## 4             19       12 0.6315789
```

, $ ( ):

```
t$lengths
## [1] 28 40 45 19 38
```

```
t$stations
## [1] 20 21 22 12 24
```

, :

```
max(t$stations)
## [1] 24
t$lengths / t$stations
## [1] 1.400000 1.904762 2.045455 1.583333 1.583333
```

colnames()

```
colnames(t)
## [1] "colors"   "lengths"  "stations" "dens"
```

, :

```
row<-data.frame("      ", 40.5, 22, 22/45)
```

, - , . ,
colnames()

```
colnames(row) <- colnames(t)
```

. , , .
, :

```
t<-rbind(t,row)
```

, :

```
colnames(t)<-c("  ","   ","    ","     ")
colnames(t)
## [1] "  "      "   "     "    "   "    "
```

:

```
t$
## [1] 28.0 40.0 45.0 19.0 38.0 40.5
t
##
## 1          28.0      20 0.7142857
## 2          40.0      21 0.5250000
## 3           45.0      22 0.4888889
## 4         19.0      12 0.6315789
## 5         38.0      24 0.6315789
## 6         40.5      22 0.4888889
```

4.3

— R. , .
, , , , — .
. .

, , :

```
d <- "                    6                "
s <- summary(t)  # summary()                    ,
```

:

```
metrolist <- list(d,t,s)
metrolist
## [[1]]
## [1] "                    6               "
##
## [[2]]
##
## 1          28.0      20 0.7142857
## 2          40.0      21 0.5250000
## 3          45.0      22 0.4888889
## 4         19.0      12 0.6315789
## 5         38.0      24 0.6315789
## 6         40.5      22 0.4888889
##
## [[3]]
##
##         :1   Min.   :19.00   Min.   :12.00   Min.   :0.4889
##        :1   1st Qu.:30.50   1st Qu.:20.25   1st Qu.:0.4979
##         :1   Median :39.00   Median :21.50   Median :0.5783
##        :1   Mean   :35.08   Mean   :20.17   Mean   :0.5800
##          :1   3rd Qu.:40.38   3rd Qu.:22.00   3rd Qu.:0.6316
##        :1   Max.   :45.00   Max.   :24.00   Max.   :0.7143
```

:

```
metrolist <- list(desc = d, table = t, summary = s)
metrolist
## $desc
## [1] "                    6               "
##
## $table
##
## 1          28.0      20 0.7142857
## 2          40.0      21 0.5250000
## 3          45.0      22 0.4888889
## 4         19.0      12 0.6315789
## 5         38.0      24 0.6315789
## 6         40.5      22 0.4888889
##
## $summary
##
##         :1   Min.   :19.00   Min.   :12.00   Min.   :0.4889
##        :1   1st Qu.:30.50   1st Qu.:20.25   1st Qu.:0.4979
##         :1   Median :39.00   Median :21.50   Median :0.5783
##        :1   Mean   :35.08   Mean   :20.17   Mean   :0.5800
##          :1   3rd Qu.:40.38   3rd Qu.:22.00   3rd Qu.:0.6316
##        :1   Max.   :45.00   Max.   :24.00   Max.   :0.7143
```

:

```
metrolist$summary
##
##         :1   Min.   :19.00   Min.   :12.00   Min.   :0.4889
##        :1   1st Qu.:30.50   1st Qu.:20.25   1st Qu.:0.4979
```

```
##          :1   Median :39.00   Median :21.50   Median :0.5783
##         :1   Mean   :35.08   Mean   :20.17   Mean   :0.5800
##           :1   3rd Qu.:40.38   3rd Qu.:22.00   3rd Qu.:0.6316
##         :1   Max.   :45.00   Max.   :24.00   Max.   :0.7143
```

summary , :

```
metrolist$summary[,3]
##
## "Min.   :12.00  " "1st Qu.:20.25  " "Median :21.50  " "Mean   :20.17  "
##
## "3rd Qu.:22.00  " "Max.   :24.00  "
```

, :

```
metrolist[[1]]
## [1] "                    6               "
metrolist[["desc"]]
## [1] "                    6               "
```

.