'

11/25/2020

Contents

Chapter 1

:

title: ,   ,
more realistic title:
c1:
c2:
c3:   (,   ,   )
c4:   :
c5:   /   (Rmd)
c6:   (git / GitHub)
c7:   Mac,   `R`? Docker.
c8:
c9:

Chapter 2

,   ,

·   ”   “

·

·

,   ,

,   ·   ·

,   ,   ·

,   ,

,   ,   ,   ·

”   “ (   )

·   ,   1)

,   2)   ·

(-   )

/   ·

·   ,

,   ·

,   ,   , (

).   ,   ,   ,

,   ,   ,

,   ·

·   ,   ,

,   ,   ·

2.1

.   (,   ,   ,   , ...),   (
,   ,   ),   (,   ,   ,   ).
,   ,
,   .   ,   ,   ,   .
,   ,   .   ,
.

25 cm,   .

,   :

- ~

- ~

- ~
,   ,
,
(   ).   ,   ,
(   ),
(   ).   ,
,   .
,   ,
.   ,   .
,   ,   ,
.   ,
,   .
,   ,   ,   .
,
.

2.2

.
,   .
,   (2:   ),
,   "   " (
3:                  ... ???).   ,
(4:   :
)
(5:   /   (Rmd)).

(　'　6:　　　　　　　　git　　(git / GitHub))

(　'　7:　　　　　Mac,　　　　R? Docker.).

Chapter 3

3.1   : R

, , Python (
: pandas, numpy, matplotlib) (: dplyr, data.table,
ggplot2). , c#
, , c# e ,
, . , Julia Scala
Python ili R
c c#. ,
.

R . R
, , ,
. R
R Rstudio
, (Rmarkdown), (bookdown), (shiny),
(blogdown), . e bookdown,
( ) , ,
Rstudio.

R.
R .
R .

R Rstudio (Rmarkdown)
Jypiter
(VScode). ,
( ) , Rstudio
. , ( )
. ,

( )   ,

.   ,

.

3.2   : git

,   git (??),

, . 				, git

e   undo/redo   ,   (   ),

,   t

,

.

git (   ).   , git

.

.   ,

. git

.

3.3   : docker

,   ,

,   ,

(big data)   .

(,   ),   (,   ),

(,   )

.   docker (??). docker

(   )   (R   )   (R   )

Rstudio   .

,   R

,

.

3.4

Table 3.1: .

		csv, MySQL
		R, git
		dplyr, ggplot2
		docker

Chapter 4

‒

4.1   (,   ,   )

.

.   ,   ,

10   R   (   ).

‒

.   , ,   ,   Downloads   :

```r
trosoci <- read_csv("~/Downloads/trosoci-moja-firma.csv")

trosoci_sumirani <- trosoci %>%
  group_by(vraboten, tip_na_trosok) %>%
  summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

write_csv(trosoci_sumirani,
          path = "~/Download/trosoci-moja-firma-sumirani.csv")
```

,   .

.   .   .

‒   .

.   Rstudio   :

```r
trosoci <- read_csv("~/Downloads/trosoci-moja-firma.csv")

trosoci_sumirani <- trosoci %>%
  group_by(vraboten, tip_na_trosok) %>%
  summarise_at("cena", "sum") %>%
```

```
  arrange(vraboten, tip_na_trosok)

write_csv(trosoci_sumirani,
          path = "~/Download/trosoci-moja-firma-sumirani.csv")
```

,   ,   .   : ”

   ?!?“   ?   :

1.   `~/Downloads/trosoci-moja-firma.csv`.
                       ?   (   )

2.   `~/Downloads/trosoci-moja-firma.csv`.
                    `Linux`   `Windows`   ,
            `Downloads`   (. `C:\Downloads`)
   `~/Downloads`)?   (   )   (   )

3.   ,   ,   -
   R: `readr`   `dplyr`.
        (   ).
                    ?   (   )   (   ).

              ,   .

500

     .

                    -     .

   :

4.   `~/Downloads/trosoci-moja-firma.csv`
                .   `Downloads`,
        .

5.   ,   `Downloads`   /
              .   ,
            .

6.   ,   `summarise_at`
                . ,   R   `dplyr`,
   `summarise_at`   .
   `dplyr`,   `summarise_at`   ,
       `dplyr`,   `summarise_at`
        .   ,   .

7.   ,
         ,   .
           ,   ,
        .   ,
        .

4.2   -

,

,

,

-                                . `C:/Data`
  `C:/Analizi`
-                             (: `C:/Data/` 2020/ /).
-         (#)   README   ,
  ,
-              (   )
-                      (
)
-            randomness
  (`set.seed`   R).

Chapter 5

5.1

,   .   (   )
vraboten, `tip_na_trosok` cena.   :

```
trosoci
```

```
## # A tibble: 30 x 3
##     vraboten    tip_na_trosok          cena
##     <chr>       <chr>                  <int>
##  1    .                         75
##  2      .                       81
##  3    .                         13
##  4       .              40
##  5        .                    89
##  6     .                        48
##  7        .          96
##  8     .                23
##  9     .                    84
## 10     .                       29
## # ... with 20 more rows
```

,   ,   :

```
trosoci %>%
  group_by(vraboten, tip_na_trosok) %>%
  summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)
```

```
## # A tibble: 12 x 3
## # Groups:    vraboten [5]
```

```
##     vraboten    tip_na_trosok              cena
##     <chr>       <chr>                     <int>
## 1     .                        31
## 2     .                            177
## 3     .                             51
## 4         .                 324
## 5         .               111
## 6          .           196
## 7          .               218
## 8          .               54
## 9         .            23
## 10        .                   48
## 11        .               127
## 12        .               116
```

5.2

(

),   .   ,   :

    1.   R

    2.

    3.

,   .

.   :

```
#
library(dplyr)
library(readr)

#               ,           :
# install.packages("dplyr")
# install.packages("readr")

trosoci <- read_csv("data/trosoci-moja-firma.csv")
```

```
## Parsed with column specification:
## cols(
##   vraboten = col_character(),
##   tip_na_trosok = col_character(),
##   cena = col_double()
## )
```

```
trosoci_sumirani <- trosoci %>%
  group_by(vraboten, tip_na_trosok) %>%
  summarise_at("cena", "sum")%>%
  arrange(vraboten, tip_na_trosok)
write_csv(trosoci_sumirani,
          path = "data/trosoci-moja-firma-sumirani.csv")
```

          ,   .
                 .   ,
          ,   ,   ,   ,
              :

```
#
library(dplyr)
library(readr)

#             ,            :
# install.packages("dplyr")
# install.packages("readr")

#     :
#
#
pateka_do_input <- NULL # "data/trosoci-moja-firma.csv"
pateka_za_output <- NULL # "data/trosoci-moja-firma-sumirani.csv"

#      :
# pateka_do_input <- "~/Downloads/trosoci-moja-firma.csv"
# pateka_do_output <- "~/Downloads/trosoci-moja-firma-sumirani.csv"

#   :
# pateka_do_input <- "C:\rabota\podatoci\trosoci\trosoci-moja-firma.csv"
# pateka_do_output <- "C:\rabota\podatoci\trosoci\trosoci-moja-firma-sumirani.csv"

#
trosoci <- read_csv(pateka_do_input)

#
trosoci_sumirani <- trosoci %>%
  group_by(vraboten, tip_na_trosok) %>%
  summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

#
write_csv(trosoci_sumirani,
          path = pateka_za_output)
```

5.3

,

           :

```r
sumiraj_trosoci <- function(trosoci, destinacija) {

  #
  trosoci <- read_csv(trosoci)

  #
  trosoci_sumirani <- trosoci %>%
    group_by(vraboten, tip_na_trosok) %>%
    summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

  #
  write_csv(trosoci_sumirani,
            path = destinacija)
}
```

   ,   (
moja-tabela.csv moja-tabela-medijani.csv)   .
              R (Python)   :

```r
#
#        `trosoci`

sumiraj_trosoci <- function(trosoci_tabela) {

  #
  trosoci <- read_csv(trosoci_tabela)

  #
  trosoci_sumirani <- trosoci %>%
    group_by(vraboten, tip_na_trosok) %>%
    summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

  #
  folder_name <- dirname(trosoci_tabela)
  base_name <- tools::file_path_sans_ext(basename(trosoci_tabela))
  new_name <- paste(base_name, "sumirani.csv", sep="-")
  destinacija <- file.path(folder_name, new_name)

  #
  write_csv(trosoci_sumirani, path = destinacija)
```

```
}
```

:

```
#       :
#                       R                                    `sumiraj_trosoci`

#
library(dplyr)
library(readr)

#              ,             :
# install.packages("dplyr")
# install.packages("readr")

#
#         `trosoci`

sumiraj_trosoci <- function(trosoci_tabela) {

  #
  trosoci <- read_csv(trosoci_tabela)

  #
  trosoci_sumirani <- trosoci %>%
    group_by(vraboten, tip_na_trosok) %>%
    summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

  #
  folder_name <- dirname(trosoci_tabela)
  base_name <- tools::file_path_sans_ext(basename(trosoci_tabela))
  new_name <- paste(base_name, "sumirani.csv", sep="-")
  destinacija <- file.path(folder_name, new_name)

  #
  write_csv(trosoci_sumirani, path = destinacija)
}
```

?   .   ,   ,
.
.

5.4   Rscript   R

TODO: (for windows see: https://stackoverflow.com/questions/3506007/runn ing-r-code-from-command-line-windows)

.   ,   R,   ,   .
                                          R,   R
      R   .   :

```
Rscript sumiraj_trosoci.R trosoci_dekemvri_2020.csv
```

                                  .   ,
              #! /usr/bin/Rscript.   ,
         ,   :

```
#! /usr/bin/Rscript

#      :
#                 R                          `sumiraj_trosoci`

#
library(dplyr)
library(readr)

#              ,            :
# install.packages("dplyr")
# install.packages("readr")

#
#       `trosoci`

sumiraj_trosoci <- function(trosoci_tabela) {

  #
  trosoci <- read_csv(trosoci_tabela)

  #
  trosoci_sumirani <- trosoci %>%
    group_by(vraboten, tip_na_trosok) %>%
    summarise_at("cena", "sum") %>%
  arrange(vraboten, tip_na_trosok)

  #
  folder_name <- dirname(trosoci_tabela)
  base_name <- tools::file_path_sans_ext(basename(trosoci_tabela))
  new_name <- paste(base_name, "sumirani.csv", sep="-")
```

```
  destinacija <- file.path(folder_name, new_name)

  #
  write_csv(trosoci_sumirani, path = destinacija)
}

#
dadeni_trosoci <- commandArgs(trailingOnly = TRUE)[[1]]

#
sumiraj_trosoci(trosoci = dadeni_trosoci)
```

,
data /sumiraj-trosoci-1.R
trosoci-moja-firma.csv.

Chapter 6

Final Words

We have finished a nice book.