

,

11/25/2020

Contents

1	:	5
2		7
2.1	8
2.2	8
3		11
3.1	: R	11
3.2	: git	12
3.3	: docker	12
3.4	12
4	-	15
4.1	(, ,)	15
4.2	-	17
5		19
5.1	19
5.2	20
5.3	22
5.4	Rscript R	24
6	Final Words	29

Chapter 1

:

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

Chapter 2

25 cm,

2.2

3: ... ???).

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

```

```

      ( )
      ,
      .
      .

```

3.2 : git

```

      , git ( ??),
e undo/redo , , git
      , t
      ,
      .
      git ( ). , git
      .
      .
      . git
      .

```

3.3 : docker

```

      ,
      ,
      (big data)
      ( , ), ( , ),
( , )
      .
      docker ( ??). docker
      (R ) (R )
      Rstudio
      .
      , R
      ,
      .

```

3.4

2.1):

Table 3.1:

		csv, MySQL
		R, git
		dplyr, ggplot2
		docker

Chapter 4

-

4.1 (, ,)

10 R ().

- Downloads :

```
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 :

```
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 ~> summarise(cena = sum(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_zo_output <- NULL # "data/trosoci-moja-firma-sumirani.csv"

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

#
# pateka_do_input <- "C:\\rabota\\podatoci\\trosoci\\trosoci-moja-firma.csv"
# pateka_zo_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_zo_output)

```

5.3

```

:
,

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)      :

#
#      `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/running-r-code-from-command-line-windows>)

```

Rscript sumiraj_trosoci.R trosoci_dekemvri_2020.csv

```

```

#!/usr/bin/Rscript

# (data/sumiraj-trosoci-2.R)

# :
# 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)

```

```

trosooci-moja-firma.csv.
data /sumiraj-trosoci-1.R
,
,
,
, docopt docstring,
. docopt/docstring
Python Perl
, assertthat,
. , docopt assertthat
commandArgs() stopifnot()
R.

```

```

#!/usr/bin/Rscript
# (data/sumiraj-trosoci-2.R)

'
ta : `vraboten`, `tip_na_trosok`, `cena`.

Usage:
sumiraj-trosoci-2.R <tabela_so_trosoci>

' -> doc

#
library(docopt)
arguments <- docopt(doc)

# CSV

```

```

assertthat::assert_that(assertthat::has_extension(arguments$tabela_so_trosoci, ext = "

#
suppressPackageStartupMessages({
  library(dplyr)
  library(readr)
  library(assertthat)
})

#
# install.packages("dplyr")
# install.packages("readr")
# install.packages(assertthat)

#
# `trosoci_tabela`

sumiraj_trosoci <- function(trosoci_tabela) {

  #
  trosoci <- read_csv(trosoci_tabela)

  assertthat::assert_that(
    inherits(trosoci, "data.frame"),
    msg = " `data.frame`." )

  assertthat::assert_that(
    all(c("vraboten", "tip_na_trosok", "cena") %in% names(trosoci)),
    msg = " : 'vraboten', 'tip_na_trosok', 'cena'." )

  assertthat::assert_that(
    is.numeric(trosoci$cena),
    msg = " `cena` ." )

  #
  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 <- arguments$tabela_so_trosoci

#
sumiraj_trosoci(trosoci = dadeni_trosoci)

```

```

#
$ Rscript sumiraj-trosoci-2.R
Error:
  ta : `vraboten`, `tip_na_trosok`, `cena`.

Usage:
  sumiraj-trosoci-2.R <tabela_so_trosoci>

Execution halted

#
$ Rscript sumiraj-trosoci-2.R trosoci-moja-firma.csv
[1] TRUE
Parsed with column specification:
cols(
  vraboten = col_character(),
  tip_na_trosok = col_character(),
  cena = col_double()
)

#
Rscript sumiraj-trosoci-2.R trosoci-moja-firma.xls
Error: File 'trosoci-moja-firma.xls' does not have extension csv
Execution halted

#
`cena`, `eur`
Rscript sumiraj-trosoci-2.R trosoci-moja-firma-eur.csv
[1] TRUE
Parsed with column specification:

```

```
cols(  
  vraboten = col_character(),  
  tip_na_trosok = col_character(),  
  eur = col_double()  
)  
Error:                               : 'vraboten', 'tip_na_trosok', 'cena'.  
Execution halted
```

Chapter 6

Final Words

We have finished a nice book.