

,

2020-12-15



# Contents



# Chapter 1

“**2:**”), “**3:**

1.3.

7

```
(
  4:
  5:
)
( 5: ' / ' (Rmd)).
git
( ' 6: )
( ' 7:
Mac, R? Docker.).
```

## 1.3

“ , (SDC)

### 1.3.1 :

title: , ,  
more realistic title:  
done: c1:  
c2:  
c3: ( , , )  
c4: ;  
c4.5:  
yet to do:  
c5: / (Rmd)  
c6: (git / GitHub)  
c7: Mac, R? Docker.  
c8:  
c9:





# Chapter 2

## 2.1 : R

```

Python (
: pandas, numpy, matplotlib) R( : 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)
Jupyter
(VScode). ( ) ,
( ) Rstudio
. , ( )

```

```

    ( )
    ,
    .
    ,
    .

```

## 2.2 : git

```

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

```

## 2.3 : docker

```

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

```

## 2.4

??):

Table 2.1:

		csv, MySQL
		R, git
		dplyr, ggplot2
		docker



# Chapter 3

## 3.1 ( , , )

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. readr dplyr. R:
   ( ).
   ? ( ) ( ).

500 , .

- :

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

5. , Downloads /
   ,

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

7. ,
   ,
   ,
   ,

```

3.2

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





# Chapter 4

## 4.1

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

## # A tibble: 30 x 3
##   vraboten    tip_na_trosok      cena
##   <chr>      <chr>          <dbl>
## 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>	<dbl>
## 1	.	31	
## 2	.		177
## 3	.		51
## 4	.		324
## 5	.		111
## 6	.	196	
## 7	.		218
## 8	.		54
## 9	.	23	
## 10	.		48
## 11	.		127
## 12	.		116

## 4.2

- (
- ), . , :
1. R
  - 2.
  - 3.

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

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

trosoci <- read_csv("data/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 = "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)

```

## 4.3

```

#
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)
}

```

?

.

.

## 4.4 Rscript

R

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

.

R,

R,

R

R

:

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

```

(
    .
    Rscript,
    (
    ),
):
# (data/sumiraj-trosoci-1.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)

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

# (data/sumiraj-trosoci-2.R)
'
ta      : `vraboten`, `tip_na_trosok`, `cena`.

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

Options:
  --help
  --version

' -> doc

#
library(docopt)
arguments <- docopt(doc, version = " 2.0\n")

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

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

```

## 4.5

•

- 
- , ,
- / —
- - ( ) ,
  - R 10 ~/Downloads Linux
  - .
  - .
  - ,
  - o , R Rscript ,
  - .

## Chapter 5

$$\{\text{rmd}\} \quad (\mathbf{R} + \text{markdown})$$

,  
 ,  
 ,  
 (t-vrednosti, p-  
 ),  
 ,  
 R  
 Rmarkdown.  
 Rmarkdown  
 ,

## 5.1 markdown

```

Markdown,      markup      HTML      LaTeX
      .
      markdown,
      R
Rstudio      Rmarkdown      R for Data Science
(      )      R      ,
      Rmarkdown („      “      ‘literal programming’).
      ,      .      ,      :
      **markdown**
[      ](https://kbroman.org/knitr_knutshell/pages/Rmarkdown.html)

```

[ ] ([https://rmarkdown.rstudio.com/authoring\\_quick\\_tour.html](https://rmarkdown.rstudio.com/authoring_quick_tour.html))  
 `Rmarkdown`.

, ja : ` 3.14 \* 2`.  
 K (chunk)  
 \*\* `R` `r`\*\*. ( - , ` ` )  
 :  $P = r^2 * \pi$   
 :  
 markdown Rmarkdown.  
 , ja : 6.28. K  
 (chunk) . R r. (  
 - ,  
 .)  
 , :  $P = r^2 * \pi$   
 HTML, PDF, Word  
 , , , .

## 5.2 , , knitr

( ) . Rmarkdown  
 ````{r}  
 ````. R knitr  
 R #.  
 knitr .  
 :  
 , :  
 rmarkdown knitr.  
 HTML, LaTeX, MS Word,  
 markup Rmd. ,  
 , , HTML LaTeX ( ???).

## 5.3

,  
 ? j

```

|---
title: "Проба"
author: "Душко долгоушко"
date: "12/14/2020"
output: html_document
---

```

```

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

```

## ## R Markdown

Ова е R Markdown документ

Кога ќе го кликнете копчето **Knit** во `Rstudio`, или извршите `rmarkdown::render()` во `R` конзола ќе се генерира документ што ја вклучува содржината и резултатит од интегрираниот R код. На пример:

```

```{r cars}
summary(cars)
```

```

## ## Вклучување графици

```

```{r pressure, echo=FALSE}
plot(pressure)
```

```

Параметарот `echo = FALSE` го додадовме за да го скриеме прикажеме `R` кодот што го прави графикот

Figure 5.1: Rmd , ,

## Проба

Душко долгоушко

12/14/2020

### R Markdown

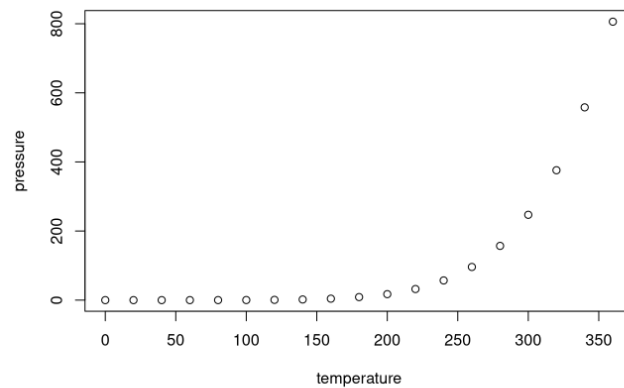
Ова е R Markdown документ

Кога ќе го кликнете копчето **Knit** во Rstudio, или извршите `rmarkdown::render()` во R конзола ќе се генерира документ што ја вклучува содржината и резултатот од интегрираниот R код. На пример:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

### Вклучување графици



Параметарот `echo = FALSE` го додадовме за да го *скриеме* прикажеме R кодот што го прави графикот

Figure 5.2: HTML

Rmd

```

Rmd
yaml

---
title:
output: html_document
params:
  grad: Tetovo
---

params R

params$grad

library(dplyr)
filtriraj_gradovi <- function(podatoci, potreben_grad) {
  podatoci %>% dplyr::filter(grad == potreben_grad)
}

#
filtriraj_gradovi(podatoci = moi_podatoci, potreben_grad = params$grad)
#

R :
rmarkdown::render(input = "mojizvestaj.Rmd", params = list("Tetovo"))

R for
HTML, PDF MS Word :
gradovi <- c("Tetovo", "Gostivar", "Debar", "Berovo", "Dojran") # ...
for ( i in gradovi) {
  message(" : ", i)
  rmarkdown::render(input = "pateka/do/mojizvestaj.Rmd", params = list(i))
}

?
izvestaj.Rmd

izvestaj.Rmd
tetovo-izvestaj.Rmd
debar-izvestaj.Rmd
skopje-izvestaj.Rmd
skopje-izvestaj-juni.Rmd

```

skopje-izvestaj-juni-specijalen-so-logo.Rmd

kichevo-izvestaj-avgust-2019.Rmd

kicevo-izvestaj-avgust.Rmd

, izvestaj.Rmd , ,

. 2-3

.

,

,

.

, kopje-izvestaj-juni-specijalen-so-logo.Rmd

.

## 5.4

rmarkdown knitr

,

:

- ( : , , )
- Rmd Shiny server ( )
- HTML prettyprint ( )
- bookdown ( )
- Rmarkdown rmarkdown::render\_site (https://rmarkdown.rstudio.com/lesson-13.html)
- flexdashboard



# Chapter 6

Rmd (REF 5). ( 4)

1:

2:

3:

6.1 1:

### 6.1.1

```
( ?) README README
```

```
strategija_1/  
  grafik_1  
    code_grafik_1.R  
    data_grafik_1.csv  
  grafik_2  
    code_grafik_2.R  
    data_grafik_2.csv  
README  
tabela_1  
  code_tabela_1.R  
  data_tabela_1.csv  
tabela_2  
  code_tabela_2.R  
  data_tabela_2.csv  
tabela_3  
  code_tabela_3.R  
  data_tabela_3.csv
```

$$\begin{matrix} 1 & 2 \\ \vdots & \vdots \\ 1 & 2, \end{matrix},$$

```
strategija_1/  
  grafik_1  
    code_grafik_1.R  
    code_tabela_1.R  
    data_grafik_1.csv  
  grafik_2  
    code_grafik_2.R  
    code_tabela_2.R  
    data_grafik_2.csv  
README  
tabela_3  
  code_tabela_3.R  
  data_tabela_3.csv
```

```
code_grafik_1.R
code_tabela_1.R
data_grafik_1.csv
```

```
data_grafik_1.csv
```

grafik\_2

```
code_grafik_2.R
code_tabela_2.R
data_grafik_2.csv
```

data\_grafik\_2.csv

# README

tabela\_3

```
code_tabela_3.R
data_tabela_3.csv
```

### 6.1.2

$$\vdots$$

### 6.1.3

 $(\quad, \quad),$ 

•

### 6.2.1

```

:

strategija_2/
  izvestai
  podatoci
  README
  skripti

R ( " "
library(mojpaket)),
:
(skripti R),
(podatoci data data-raw), README
, .
?
,
:
1. R (global environment, .GlobalEnv),
2. podatoci
3. ( , , HTML/PDF ) izvestai
,
,

```

### 6.2.2

```

,
.
R , .
:
```

```

library(devtools)
library(usethis)
devtools::create(path = "mojpaket")

podatoci1 <- read.csv("~/Desktop/moi-podatoci.csv")
usethis::use_data("podatoci1")

```

```

:
```

```

mojpaket/
DESCRIPTION
NAMESPACE
data
R

```

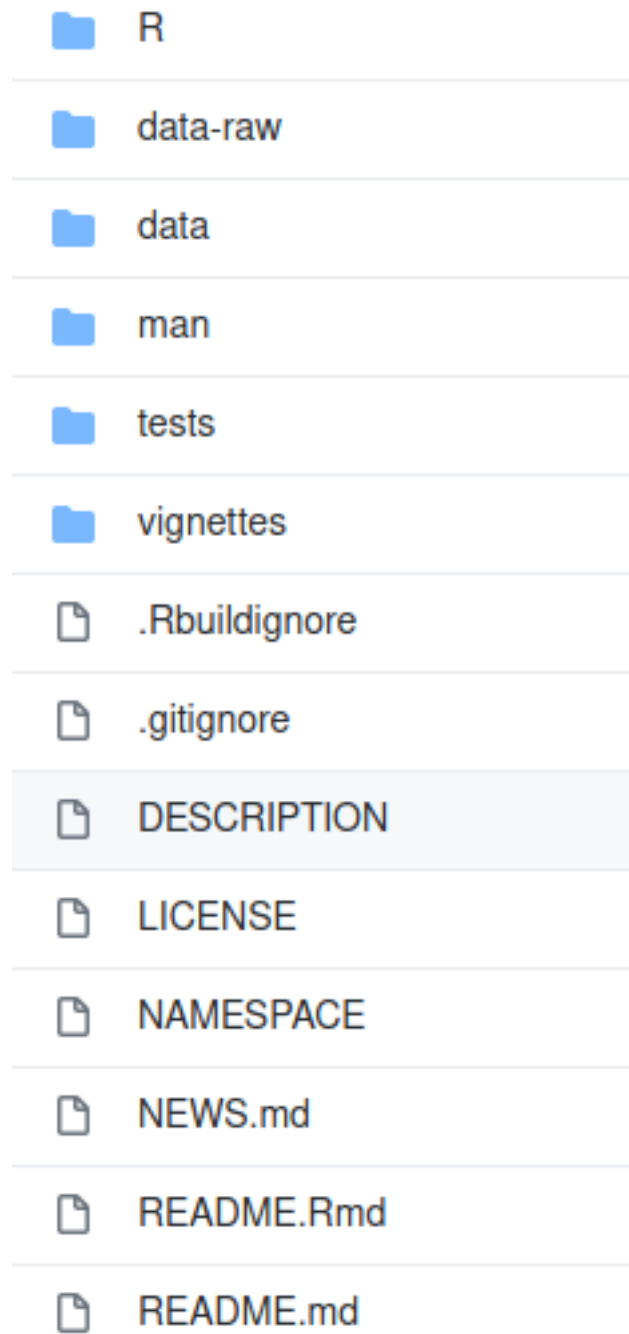


Figure 6.1: R ( )

```

, , , , R ((?)).
:
/ /
: prodigenr, makeProject, ProjectTemplate,
fertile, goodpractice, (
, ), , , prodigenr
R (
R
library(prodigenr)
setup_project("data/mojproekt2")

, :
mojproekt2/
data
  README.md
  DESCRIPTION
  doc
  README.md
  mojproekt2.Rproj
  R
  fetch_data.R
  README.md
  setup.R
  README.md
  TODO.md
, setup.R /
  fetch_data.R
data/
, DESCRIPTION ( - R ).
( library(help="prodigenr"))

```

### 6.2.3

```

R . R
, , (unit
tests)

```

### 6.2.4

```

, ,

```

, , ,  
, .



# Chapter 7

## (git / GitHub)

Git. ( Git ) Github Gitlab.  
 , Git  
 , , . track  
changes MS Word.  
 , Git .  
 . https://git-scm.com/book/en/v2  
Github HMTL.



## Chapter 8

not too much about git (of course that is not the point anyway)



## Chapter 9

just about creating a github repo where you can push rmd in main and html in gh-pages thus making your reserarch public on the web mostly working thourgh r studio exept for creating the account on gh.

9.0.1 : <https://git-scm.com/book/en/v2>