

Uvod u R i RStudio

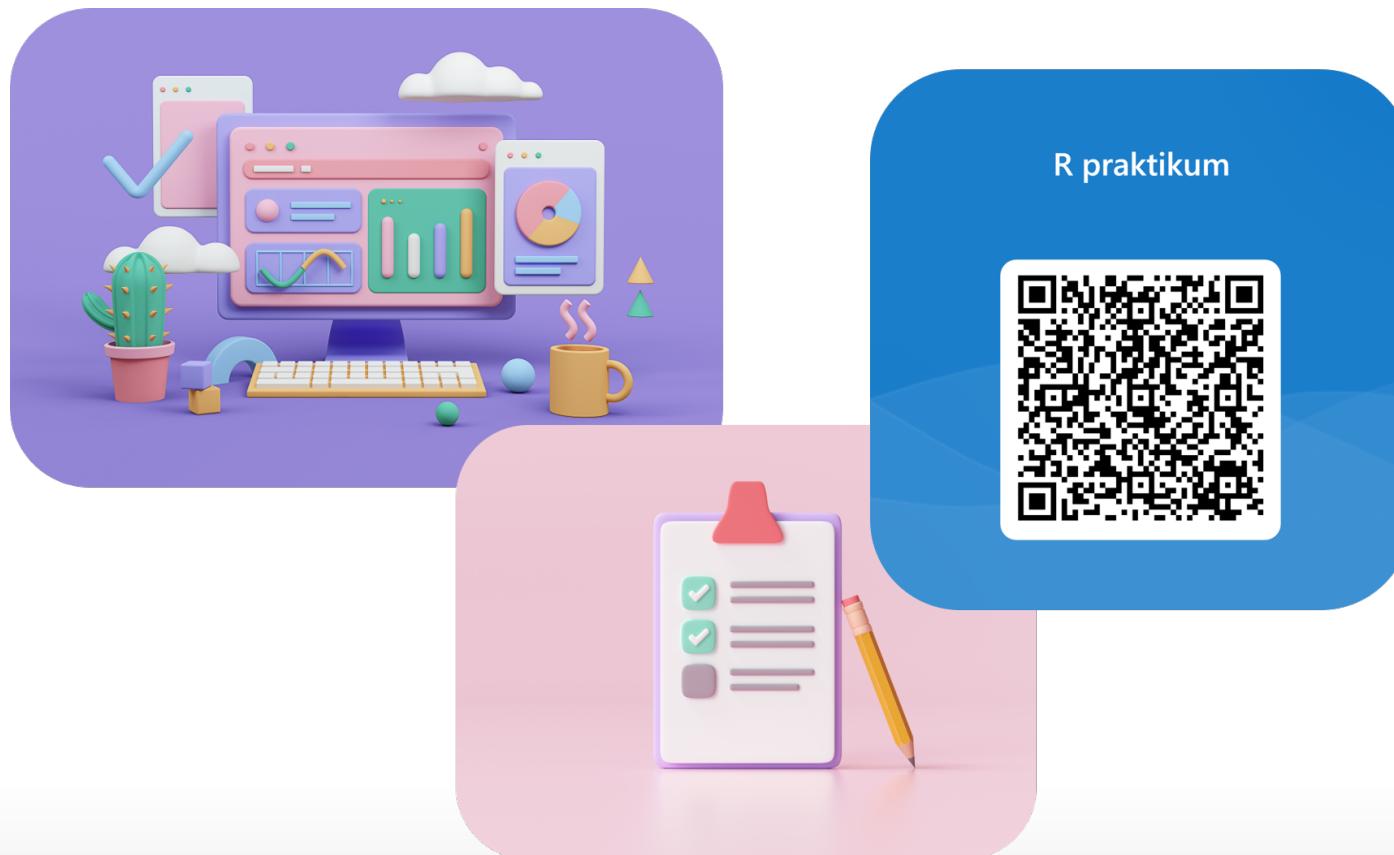
Analize podataka u biološkim istraživanjima

dr. sc. Lucija Kanjer, e-mail: lucija.kanjer@biol.pmf.hr

2025/26

Anketa upoznavanja

- Otvorite poveznicu: <https://forms.office.com/e/FPkBKCs02J> ili skenirajte QR kod
- prijavite se sa svojim AAIEDU identitetom
- ispunite kratku anketu



Teme praktikuma

- Uvod i unos podataka u R
- Rad s podacima
- Grafički prikazi
- Razdiobe i deskriptivna statistika
- Parametrijska statistika
- Neparametrijska statistika
- Statistika kategoričkih podataka
- Multivariatna statistika

Raspored po tjednima

Tjedan	Tema vježbe
06.10.-12.10	Uvod i unos podataka u R
13.10.-19.10	Rad s podacima
20.10.-26.10.	Grafički prikazi 1
27.10.-02.11.	Grafički prikazi 2 i razdiobe
03.11.-09.11.	Deskriptivna statistika i vježba za kolokvij 1
10.11.-16.11.	Kolokvij 1
17.11.-23.11.	<i>Jesenski ispitni rokovi</i>
24.11.-30.11.	Parametrijska statistika
01.12.-07.12.	Neparametrijska statistika i statistika kategoričkih podataka
08.12.-14.12	LMs i GLMs
15.12.-21.12.	Multivarijatna statistika 1
22.12.-28.12.	<i>Rezervni termin (pon i uto)</i>
29.12.-04.01.	<i>Božićni praznici</i>
05.01.-11.01.	<i>Rezervni termin (sri i čet)</i>
12.01.-18.01.	Multivarijatna statistika 2
19.01.-25.01.	Vježba za kolokvij 2
26.01.-01.02.	Kolokvij 2

Sadržaj današnje vježbe

- Upoznavanje s RStudio sučeljem
- Izvršavanje naredbi iz skripte
- definiranje i ispis objekata
- pisanje komentara
- osnovne matematičke operacije
- postavljanje radnog direktorija
- traženje pomoći
- instalacija i učitavanje paketa
- primjer izrade grafičkog prikaza

Ishodi učenja praktikuma

- Upoznati se s R okruženejm
- Premostiti strah od „programiranja“
- Proći „energiju aktivacije“ koja je potrebna za korištenje R-a i rada na vlastitim podacima i projektima
- Moći izraditi vlastite analize pomoću svih danas dostupnih alata

Što nije ishod ovog praktikuma?

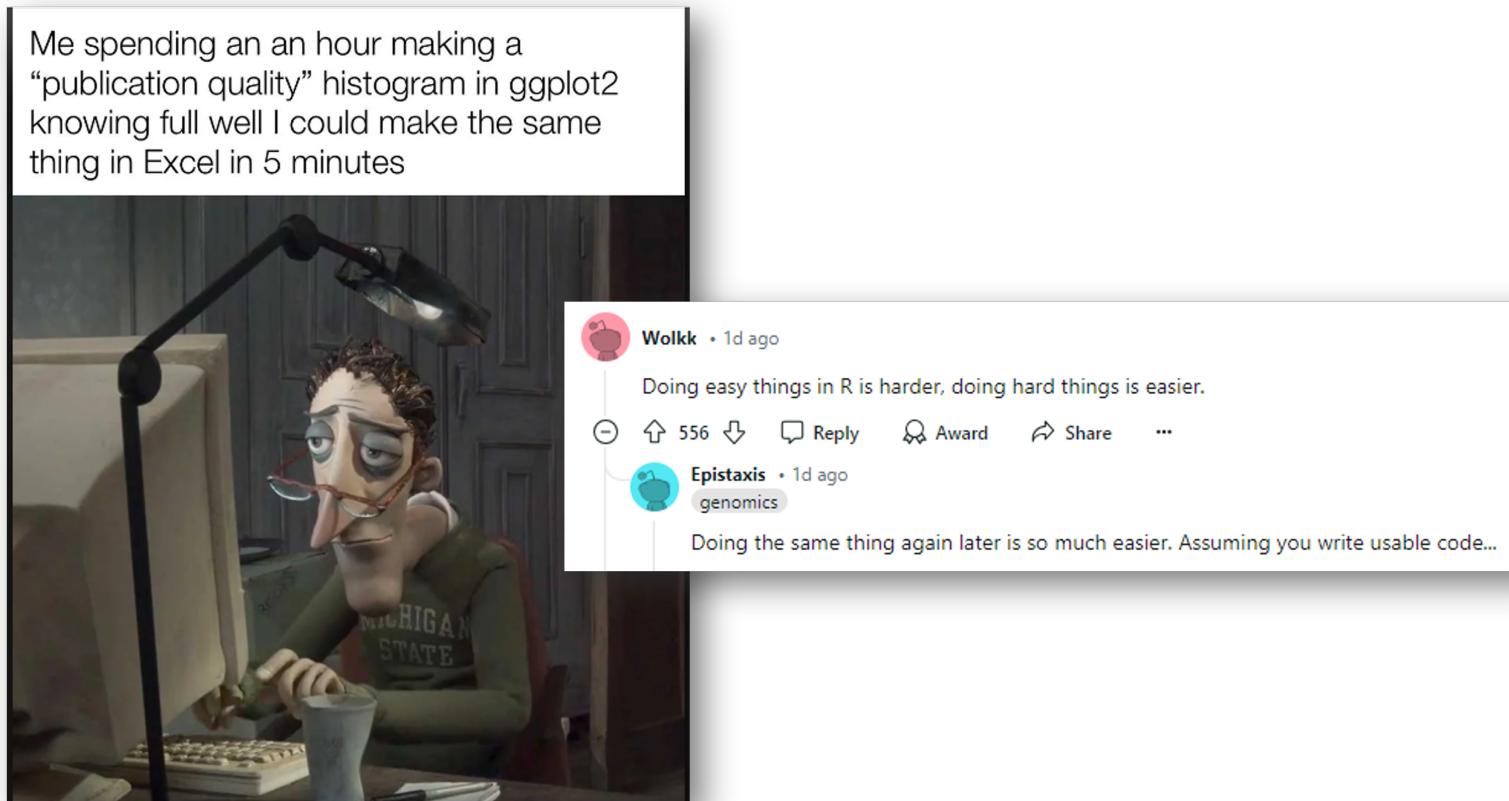
- Sveobuhvatno znanje o R-u
- Sveobuhvatno znanje o svim statističkim metodama
- Tečaj programiranja
- Učenje napamet naredbi R-a

Prednosti R-a

1. Besplatan softver otvorenog koda.
2. Velika zajednica korisnika i podrška.
3. Prilagođen akademskoj zajednici (i biologima).
4. Fleksibilnost i prilagodljivost - veliki broj paketa za specijalizirane analize.
5. Napredne statističke i analitičke metode.
6. Vizualizacija podataka.
7. Reproducibilnost.

Mane R-a u odnosu na *point and click* programe

- Strma krivulja učenja.
- Manje intuitivno sučelje.



Što nam treba za rad?

R - statistički programski jezik



RStudio - multifunkcionalno sučelje za rad s programskim jezikom R, engl. *integrated development environment* (IDE), mogu se koristiti i drugi programi kao što su VS Code ili Positron



Instalacija R-a

R se može preuzeti sa CRAN-a (Comprehensive R archive network) <https://cran.r-project.org/>

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

- [Download R for Linux \(Debian, Fedora/Redhat, Ubuntu\)](#)
- [Download R for macOS](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you can also download R from your distribution's repositories.

R for Windows

Subdirectories:

- [base](#) Binaries for base distribution. This is what you want to [install R for the first time](#).
- [contrib](#) Binaries of contributed CRAN packages (for R >= 4.0.x).
- [old contrib](#) Binaries of contributed CRAN packages for outdated versions of R (for R < 4.0.x).
- [Rtools](#) Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

Please do not submit binaries to CRAN. Package developers might want to contact Uwe Ligges directly in case of questions / suggestions related to Windows binaries.

You may also want to read the [R FAQ](#) and [R for Windows FAQ](#).

Note: CRAN does some checks on these binaries for viruses, but cannot give guarantees. Use the normal precautions with downloaded executables.

Instalacija RStudio-a

RStudio je dostupan za preuzimanje na Posit web stranici:

- <https://posit.co/download/rstudio-desktop/>

The screenshot shows two main sections: '1: Install R' and '2: Install RStudio'.

1: Install R
RStudio requires R 3.6.0+. Choose a version of R that matches your computer's operating system.
R is not a Posit product. By clicking on the link below to download and install R, you are leaving the Posit website. Posit disclaims any obligations and all liability with respect to R and the R website.
[DOWNLOAD AND INSTALL R](#)

2: Install RStudio
[DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS](#)
Size: 265.55 MB | [SHA-256: 513216FE](#) | Version: 2024.09.0+375 | Released: 2024-09-23

RStudio - sučelje

The screenshot shows the RStudio interface with four main panels labeled A, B, C, and D.

- Panel A:** Script Editor (top-left). It displays an R script titled "map.R" with code related to reading metadata and creating a map plot. A large black square with a white letter "A" is overlaid on the left side of the panel.
- Panel B:** Environment (top-right). It shows the global environment with objects like "carapace_samples", "metadata", "values", and "countries". A large black square with a white letter "B" is overlaid on the right side of the panel.
- Panel C:** Console (bottom-left). It shows the output of running the R script, listing various locations like Susak, Mali Losinj, and Vis. A large black square with a white letter "C" is overlaid on the left side of the panel.
- Panel D:** Plots (bottom-right). It is currently empty, showing a blank white area. A large black square with a white letter "D" is overlaid on the right side of the panel.

RStudio - sučelje

- Panel A: **Skripta** - dio u koji pišemo naredbe i komentare koje možemo izvršavati i spremiti za ponovnu upotrebu.
- Panel B: **Konzola** (engl. console) - dio u kojemo se naredbe izvršavaju (interaktivno okruženje), ali se ne može spremiti za ponovnu upotrebu; također i dio za terminal, render i pozadinske poslove.
- Panel C: **Okoliš** (engl. environment) - dio u kojem vidimo stvorene R objekte; također i dio za povijest, veze; git sinkronizaciju i tutoriale.
- Panel D: **plots** - dio za vizualizaciju grafova; **packages** - dio za vizualizaciju i učitavanje dostupnih paketa; **help** - u kojem se prikažu tražene informacije o paketima i naredbama; **files** - vizualizacija trenutnog radnog direktorija (nije za postavljanje novog radnog direktorije).

Materijali

- Materijali vježbi dostupni na GitHub repozitoriju “APUBI_2026_26” na linku:
- https://github.com/lucijakanjer/APUBI_2025_26

The screenshot shows a GitHub repository page for "APBI_2024". The repository is public and was created by "lucijakanjer". It contains one branch ("main") and several files:

- 01_Uvod_prezentacija.pdf
- 01_Uvod_skripta.R
- README.md
- zivotinje.txt

The README file contains the following text:

APBI_2024

Praktične vježbe za kolegij "Analiza podataka u biološkim istraživanjima" u akademskoj godini 2024/25

On the right side of the page, there is an "About" section with the following details:

- Praktične vježbe za kolegij "Analiza podataka u biološkim istraživanjima" u akademskoj godini 2024/25
- Readme
- Activity
- 0 stars
- 1 watching
- 0 forks

There are also sections for "Releases" and "Packages", both of which currently have no content.

Rad u R-u i RStudio-u

- Otvorite RStudio i krenimo s prvim naredbama!
- Izvršavanje naredbi: **Ctrl + Enter** u skripti.

```
# Prva naredba: ispis teksta iz skripte u konzolu
print("Hello world!")
```

```
## [1] "Hello world!"
```

Komentari

- Sve što se piše iza znaka "#" u skripti, predstavlja **komentar** i neće biti izvršeno kao naredba.
- Komentari se pišu u skripti kako bi naš kod bio pregledan i reproducibilan.

```
# Ovo je primjer komentara cijelog reda  
print("Dobar dan svijete!") # ovo je komentar nakon naredbe
```

```
## [1] "Dobar dan svijete!"
```

Definiranje varijabli

- Varijable su pojedinačne vrijednosti.
- U R-u, svaka varijabla može pohranjivati brojeve, tekstualne vrijednosti ili logičke vrijednosti.
- Definiraju se pomoći znakova "<-" ili "="

```
# Definiranje objekata (varijabli)
a <- 5
b = 3
ime <- "Lucija"
```

- prvi način je mnogo češće korišten u praksi jer se znak "=" koristi za i u druge svrhe

Ispis

Ispis tj. prikaz R objekata, može se izvršiti na više načina:

```
# Različiti načina ispisa  
print(a) # ispis u konzoli
```

```
## [1] 5
```

```
a # ispis u konzoli
```

```
## [1] 5
```

```
print(a + b) # ispis rezultata u konzoli
```

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

```
cat("Ja sam", ime) # ispis više objekata i karaktera odjednom
```

```
## Ja sam Lucija
```

Razlika između `cat()` i `print()`

Razlike između `print()` i `cat()` naredbi u R-u nalaze se u načinu na koji prikazuju izlaz i koriste se za različite svrhe:

- `print()` se koristi za ispisivanje objekata (vektora, matrica, data frameova itd.) u R-u. Ovo je najčešća funkcija za ispis rezultata.
- `cat()` se koristi za konkatenaciju (spajanje) i ispis teksta. Idealan je za ispisivanje nizova karaktera i jednostavnih poruka.

Zadatak

Napravite varijablu "visina" te u nju spremite vrijednost svoje visine, te varijablu "godina" te u nju spremite vrijednost svoje godine rođenja. Ispišite svaku varijablu na različiti način.

```
# Primjer: Definirajmo varijablu svoje visine i godine rođenja  
visina <- 170  
godina <- 1995  
# Ispišimo ove varijable!  
print(visina)
```

```
## [1] 170
```

```
cat("Ovo je moja godina:", godina)
```

```
## Ovo je moja godina: 1995
```

Rad s osnovnim matematičkim operacijama

```
# Zbroj  
zbroj <- 5 + 3  
print(zbroj)
```

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

```
# Umnožak  
umnožak <- 4 * 7  
print(umnožak)
```

```
## [1] 28
```

Zadatak

Na isti način kao za zbroj i umnožak izračunajmo i ispišimo rezultate oduzimanja(razliku) i dijeljenja brojeva (kvocijent).

Razlika i kvocijent

```
# Razlika  
razlika <- 10 - 4  
print(razlika)
```

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

```
# Kvocijent  
kvocijent <- 38 / 7  
print(kvocijent)
```

```
## [1] 5.428571
```

Ručni unos objekata u R-u

Varijable

```
### Kreiranje objekata u R-u - ručni unos
# 1. Varijable
x <- 3.14          # numeric - brojčana vrijednost
y <- "DNA"         # character - tekst (mora ići u navodnike)
z <- TRUE          # Logical - Logička vrijednost (TRUE/FALSE)
```

Provjera tipa i struktire podataka

```
# Provjera tipa podataka
```

```
class(x) # prikazuje samo tip varijable
```

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

```
str(x) # struktura: prikazuje tip varijable i što ona sadrži
```

```
## num 3.14
```

Provjerite tip podataka varijabli y i z!

```
# Provjerite tip podataka varijabli y i z!
```

```
class(y)
```

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

```
class(z)
```

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

Vektori

Vektor je niz elemenata iste vrste (npr. svi brojevi ili svi tekstovi).

```
# 2. Vektori  
geni <- c("BRCA1", "TP53", "MYC") # tekstualni vektor gena  
print(geni)
```

```
## [1] "BRCA1" "TP53"  "MYC"
```

```
duljine <- c(1863, 1179, 1584)      # numerički vektor duljina gena u parovima baza (bp)  
print(duljine)
```

```
## [1] 1863 1179 1584
```

Matrice

Matrica je 2D struktura podataka iste vrste (redovi x stupci).

```
# 3. Matrice - skup vektora istog tipa
# Definiranje matrice
matrix_mikrobi <- matrix(c(85, 47, 37, 10, 65, 50, 28, 8),
                           nrow = 4, # postavljanje broja redaka u matrici
                           byrow = TRUE) # definiranje popunjavanja matrice po retcima

# Dodavanje naziva redaka i stupaca
rownames(matrix_mikrobi) <- c("Bakterija1", "Bakterija2", "Gljivica1", "Gljivica2")
colnames(matrix_mikrobi) <- c("CFU_kuhinja", "CFU_kupaonica")
```

```
print(matrix_mikrobi)
```

	CFU_kuhinja	CFU_kupaonica
## Bakterija1	85	47
## Bakterija2	37	10
## Gljivica1	65	50
## Gljivica2	28	8

Podatkovni okvir (*data frame*)

Data frame je struktura podataka koja može sadržavati vektore varijabli različitih vrsta (brojevi, tekst, logičke vrijednosti).

```
# 4. Podatkovni okvir (data frame) - skup vektora različitih tipova
df_mikrobi <- data.frame(
  CFU = c(85, 47, 37, 10, 65, 50, 28, 8), # numerički vektor
  mikrorganizam = c("Bakterija", "Bakterija", "Gljivica", "Gljivica"), # tekstualni vektor
  lokacija = c("kuhinja", "kupaonica", "kuhinja", "kupaonica") # tekstualni vektor
)
```

```
# Uvid u data frame  
print(df_mikrobi) # ispis u konzoli
```

```
##   CFU mikrorganizam lokacija  
## 1  85    Bakterija  kuhinja  
## 2  47    Bakterija kupaonica  
## 3  37    Gljivica  kuhinja  
## 4  10    Gljivica kupaonica  
## 5  65    Bakterija  kuhinja  
## 6  50    Bakterija kupaonica  
## 7  28    Gljivica  kuhinja  
## 8   8    Gljivica kupaonica
```

```
View(df_mikrobi) # vizualizacija tablice u novom listu
```

The screenshot shows a data visualization interface with a table of microbiological data. The table has columns: CFU, mikrorganizam, and lokacija. The data consists of 8 rows, each numbered from 1 to 8. Rows 1, 2, 5, and 6 contain 'Bakterija' in the 'mikrorganizam' column, while rows 3, 4, 7, and 8 contain 'Gljivica'. Rows 1, 3, 5, and 7 are located in 'kuhinja', while rows 2, 4, 6, and 8 are located in 'kupaonica'. The 'CFU' values range from 8 to 85. The interface includes navigation buttons (back, forward, search), a filter icon, and a search bar.

	CFU	mikrorganizam	lokacija
1	85	Bakterija	kuhinja
2	47	Bakterija	kupaonica
3	37	Gljivica	kuhinja
4	10	Gljivica	kupaonica
5	65	Bakterija	kuhinja
6	50	Bakterija	kupaonica
7	28	Gljivica	kuhinja
8	8	Gljivica	kupaonica

Showing 1 to 8 of 8 entries, 3 total columns

Rad s datotakama

Radni direktorij, engl. working directory

```
# Provjeravanje radnog direktorija (working directory)
getwd()
```

```
## [1] "C:/Users/lucij/Documents/APUBI/01_Uvod"
```

```
# Postavljanje radnog direktorija
```

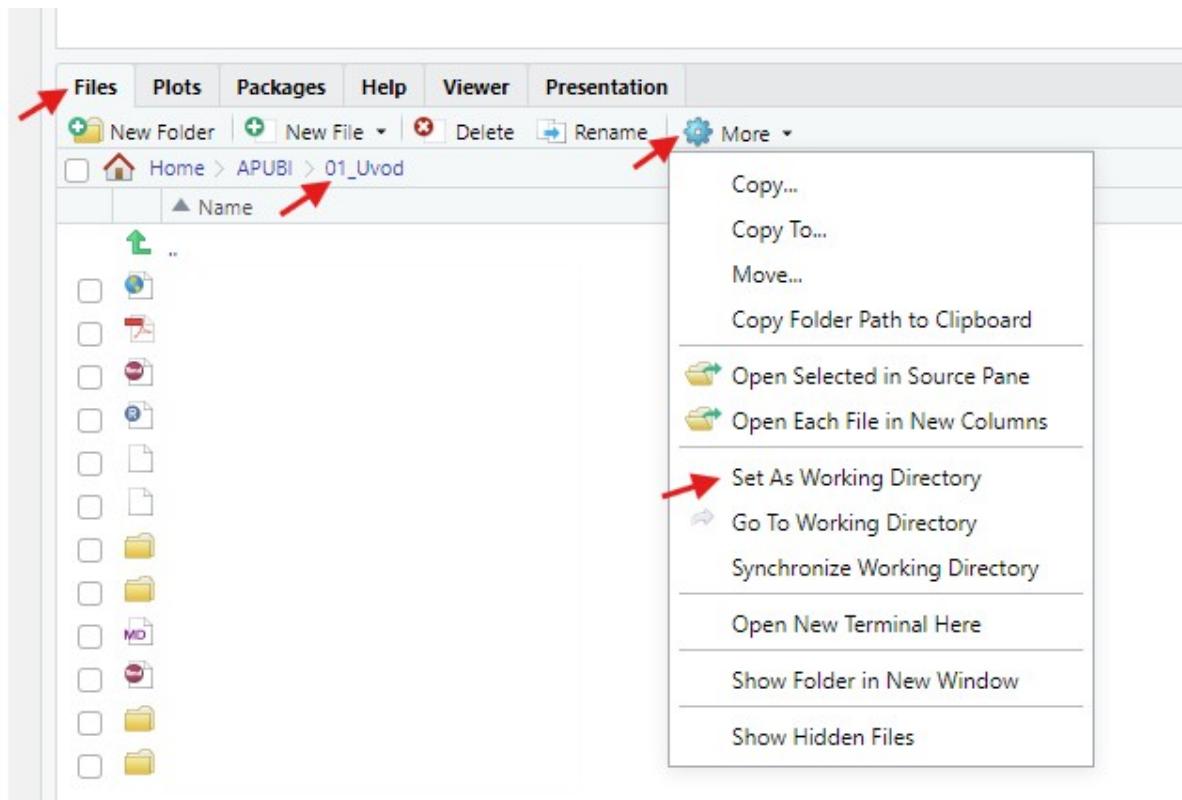
```
setwd("C:/Users/lucij/Documents/APUBI/01_Uvod/zivotinje") # moj Laptop
#setwd("C:/Users/Hrvoje/Documents/APUBI/01_Uvod/zivotinje") # poslovno računalo
```

```
# Iznad je još jedan primjer korištenja komentara:
```

```
# za naredbe koje ne želimo da se trenutno izvršavaju.
```

Postavljanje radnog direktorija

- alternativni način postavljanja radnog direktorija klikanjem
- Files* -> ući u mapu koju želimo postaviti kao radni direktorij -> *More* -> *Set As Working Directory*



Traženje pomoći

- `help(setwd)`
- `?setwd`
- `??setwd`

Učitavanje datoteke s podacima

```
# Učitavanje datoteke s podacima u objekt "zivotinje"  
zivotinje <- read.table("C:/Users/lucij/Documents/APUBI/01_Uvod/zivotinje/zivotinje.txt",  
                         header = TRUE)
```

Kako pogledati svoje podatke?

```
print(zivotinje) # ispis cijele tablice u konzoli
```

```
##      uzorak    vrsta visina_cm masa_g skupina
## 1        1     Miš       10     25 Sisavci
## 2        2     Miš       12     30 Sisavci
## 3        3     Miš        9     22 Sisavci
## 4        4     Miš       11     28 Sisavci
## 5        5     Miš       13     35 Sisavci
## 6        6   Golub       35    300   Ptice
## 7        7   Golub       32    280   Ptice
## 8        8   Golub       36    320   Ptice
## 9        9   Golub       34    310   Ptice
## 10      10   Golub       33    290   Ptice
## 11      11 Vrabac       15     24   Ptice
## 12      12 Vrabac       14     22   Ptice
## 13      13 Vrabac       16     26   Ptice
## 14      14 Vrabac       15     23   Ptice
## 15      15 Vrabac       13     21   Ptice
## 16      16 Vjeverica     20    400 Sisavci
## 17      17 Vjeverica     22    420 Sisavci
## 18      18 Vjeverica     21    410 Sisavci
## 19      19 Vjeverica     19    380 Sisavci
```

```
# Vizualizacija podataka - kratki oblik  
head(zivotinje) # prikaz "glave" tablice
```

```
##   uzorak vrsta visina_cm masa_g skupina  
## 1     1 Miš        10    25 Sisavci  
## 2     2 Miš        12    30 Sisavci  
## 3     3 Miš         9    22 Sisavci  
## 4     4 Miš        11    28 Sisavci  
## 5     5 Miš        13    35 Sisavci  
## 6     6 Golub      35   300 Ptice
```

View(zivotinje) # prikaz u novom tabu pored skripte

	uzorak	vrsta	visina_cm	masa_g	skupina
1	1	Miš	10	25	Sisavci
2	2	Miš	12	30	Sisavci
3	3	Miš	9	22	Sisavci
4	4	Miš	11	28	Sisavci
5	5	Miš	13	35	Sisavci
6	6	Golub	35	300	Ptice
7	7	Golub	32	280	Ptice
8	8	Golub	36	320	Ptice
9	9	Golub	34	310	Ptice
10	10	Golub	33	290	Ptice
11	11	Vrabac	15	24	Ptice

Showing 1 to 12 of 20 entries, 5 total columns

Paketi

- Instalacijom dodatnih paketa moguće je prilagoditi svoje analize, a izbora je stvarno puno!
- Npr. Za vizualizaciju podataka najčešće se korsiti dodatni paket **ggplot2** koji nije uključen u temeljni R (*base R*) te ga je potrebno instalirati.
- Instalacija paketa izvršava se samo jednom, a učitavanje paketa mora se izvršiti prilikom svakog otvaranja RStudio-a, tj. početkom svake nove sesije.

Paketi

```
# Za dodatne analize moramo instalirati dodatne R pakete  
# Instalacija paketa - samo jednom
```

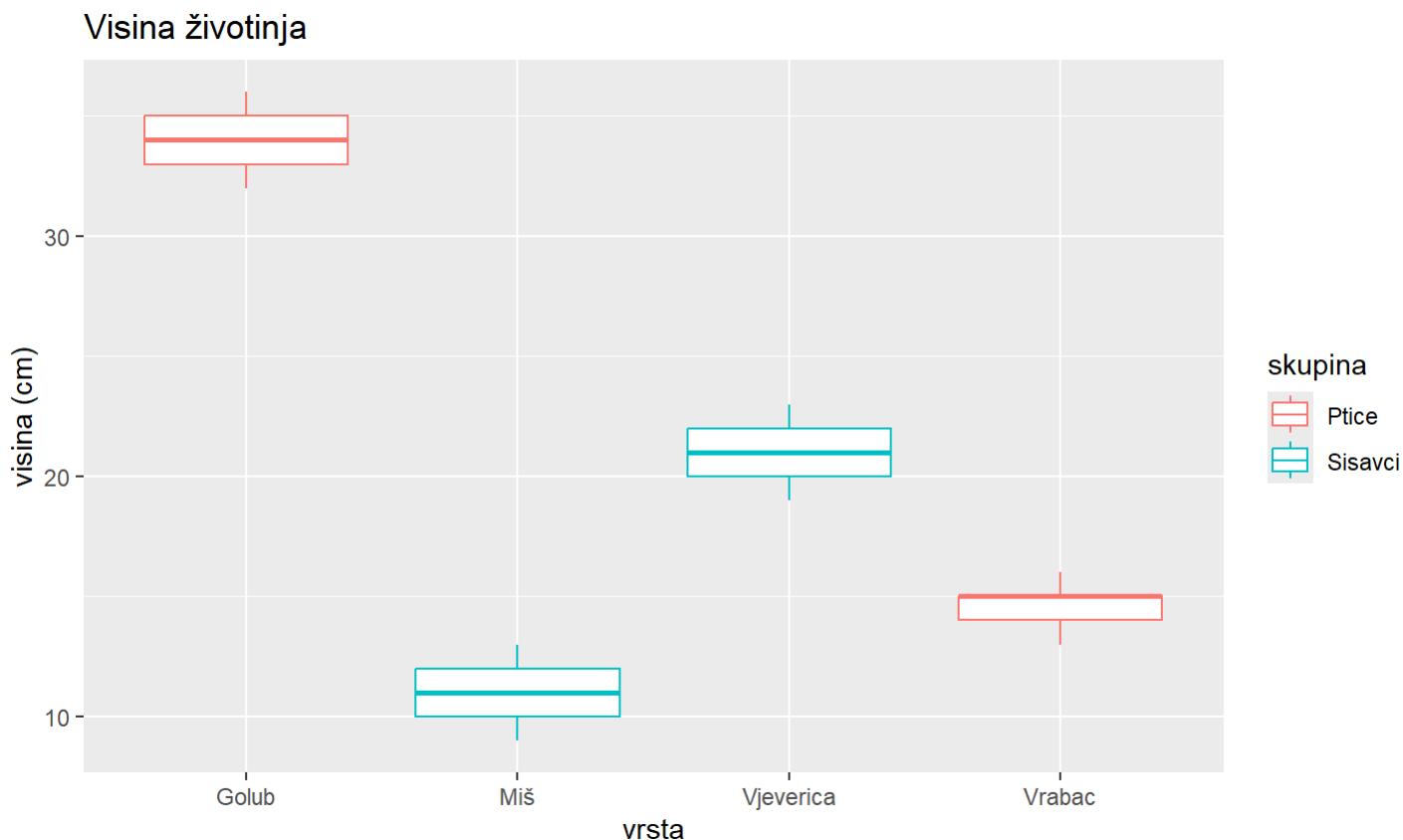
```
install.packages("ggplot2")
```

```
# Učitavanje paketa u radno okruženje - pri svakom pokretanju R-a  
library(ggplot2)
```

Primjer izrade grafa

```
# izrada objekta "graf_visina"
graf_visina <- ggplot(data = zivotinje, aes(x = vrsta, y = visina_cm)) +
  geom_boxplot(aes(color = skupina)) +
  labs(title = "Visina životinja",
       x = "vrsta",
       y = "visina (cm)")
```

```
# ispis objekta "graf_visina"  
graf_visina
```



Izmjenite gornji graf tako da se umjesto visine prikazuju podaci za masu!

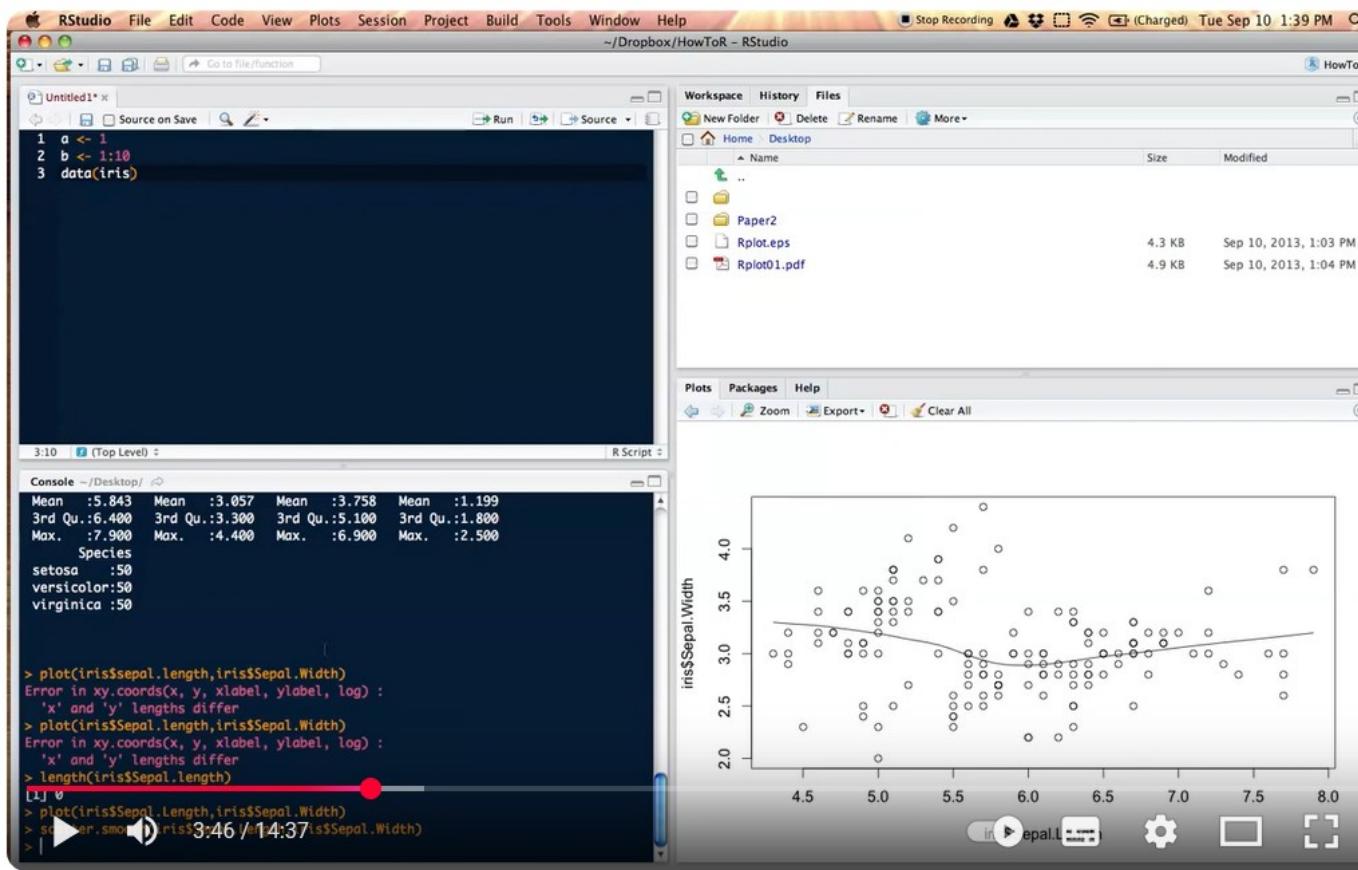
Pazite da izmjenite naslov grafa i nazive osi da odgovaraju novom prikazu.

```
# izrada objekta "graf_masa"  
graf_masa <- ggplot(data = zivotinje, aes(x = vrsta, y = masa_g)) +  
  geom_boxplot(aes(color = skupina)) +  
  labs(title = "Masa životinja",  
       x = "vrsta",  
       y = "masa (g)")
```

Dodatni izvori

How To R: Getting Started with R and RStudio

<https://www.youtube.com/watch?v=lVKMsaWju8w>



Getting started with R and RStudio



How To R

8.52K subscribers

Subscribe

3.8K



Share

Clip

...

45/47

Samostalni zadaci

1. Kreirajte 3 varijable za visine 3 različite životinje.
2. Izračunajte i ispišite umnožak visine životinja koje ste upisali.
3. Napravite novu mapu naziva "biljke" te ju postavite kao novi radni direktorij.
4. Kreirajte 2 varijable, u jednu spremite ime vrste po vašem izboru, a u drugu broj jedinki te vrste.
5. Ispište tip objekta za novokreirane varijable.
6. Kreirajte 2 vektora, u jednom pohranite tekstove, a u drugom brojeve.
7. Kreirajte data frame s biološkim podacima po vašem izboru, mora sadržavati 2 numeričke i 2 kategoričke varijable.
8. Ispišite strukturu novokreiranog data frame-a.

Kviz

Otvorite poveznicu: <https://forms.office.com/e/cG5bh8WvPk>

ili skenirajte QR kod i ispunite kratki kviz danas naučenoga!

