

Univerzitet u Beogradu – Elektrotehnički fakultet (ETF)

Katedra za signale i sisteme



Uvod u R i R Studio

Dr Nadica Miljković, vanredna profesorka
kabinet 68, nadica.miljkovic@etf.bg.ac.rs

Instalacija

- Sa CRAN-a je potrebno instalirati programski jezik R <https://cran.r-project.org/> (*Download R for Windows / Base link / Download*).
 - Preporuka: koristiti podrazumevane instalacione parametre.
 - Proveriti da li je instalacija koju ste preuzeli odgovarajuća za Vaš operativni sistem!
- Potom je potrebno instalirati R Studio (IDE – *Integrated Development Environment*). Ovaj IDE se može instalirati samo ako već postoji instaliran R na računaru <https://posit.co/downloads/> (Desktop opcija).
 - Napomena: veb sajt bi trebalo da prepozna koji operativni sistem se nalazi na računaru i da preporuči instalacionu fasciklu/folder koji je potrebno preuzeti.
 - U novijim verzijama, u ponudi je zajednička instalacija R-a i R Studio okruženja.

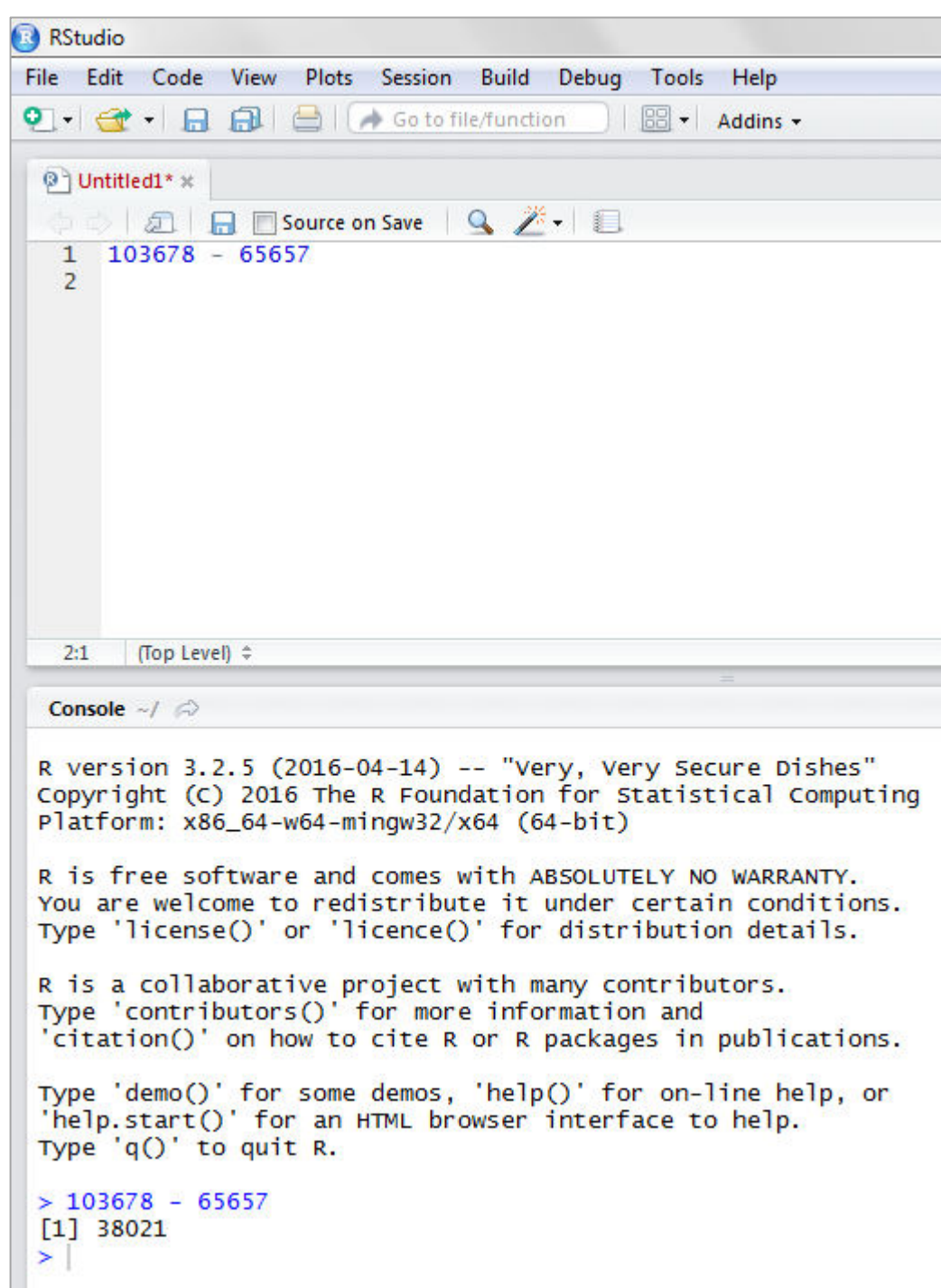


R Studio IDE

The image shows the RStudio IDE interface with four panels highlighted by orange text labels:

- SOURCE - IZVORNI**: The top-left panel shows the source editor with R code. The code includes a file path, a comment, a calculation, and a package installation command.
- ENV./HIS. - OKRUŽENJE**: The top-right panel shows the Environment, History, and Connections pane. The Environment pane is currently empty.
- VIEW/HELP... - PREGLEDAČ**: The bottom-right panel shows the Files, Plots, Packages, Help, and Viewer pane. The Help pane is currently empty.
- CONSOLE - KOMANDNI**: The bottom-left panel shows the Console and Terminal pane. The Console displays the output of the 'contributors()' function, including instructions on how to use 'demo()', 'help()', and 'help.start()'.

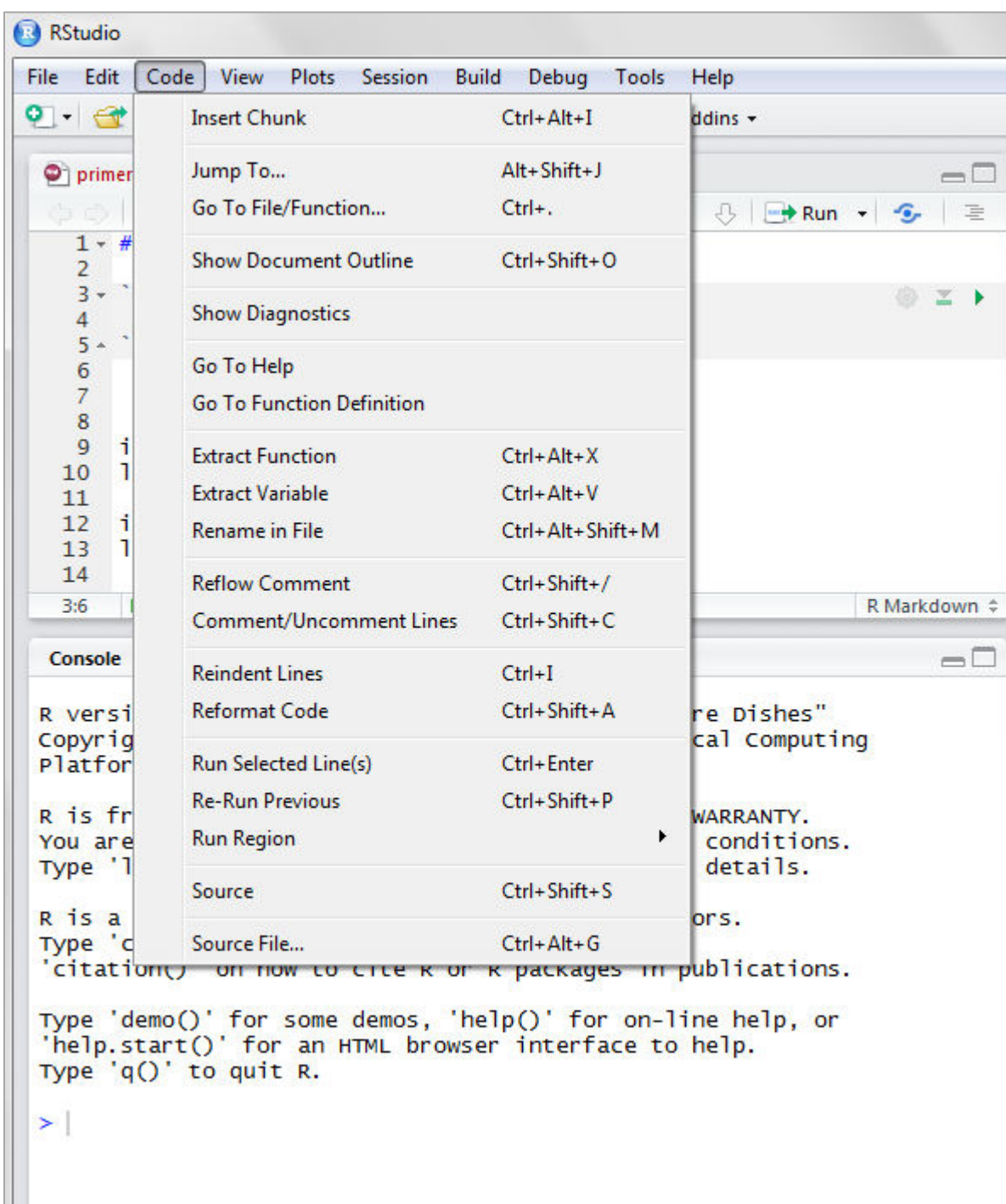
Da počnemo ...



The image shows a screenshot of the RStudio software interface. At the top, there is a menu bar with options: File, Edit, Code, View, Plots, Session, Build, Debug, Tools, and Help. Below the menu bar is a toolbar with icons for file operations and a search bar labeled 'Go to file/function'. The main workspace contains a script editor with a file named 'Untitled1*'. The script contains two lines of code: `1 103678 - 65657` and `2`. Below the script editor is a console window. The console displays the R version information: 'R version 3.2.5 (2016-04-14) -- "Very, Very Secure Dishes" Copyright (C) 2016 The R Foundation for Statistical Computing Platform: x86_64-w64-mingw32/x64 (64-bit)'. It also shows the license information and instructions for using R. At the bottom of the console, the output of the script is shown: `> 103678 - 65657` followed by `[1] 38021` and a prompt `>`.

R Studio

- Podeljen je na prozore (konzola i skripta):
 - Komandni prozor ili konzola služe za izvršavanje komandi.
 - Na slici je dat primer pokretanja komande (operacija oduzimanja) u skriptu.
 - Takođe, komanda se može zadati i u konzoli.
 - Rezultat se uvek ispisuje u konzoli.
- Brojevi koji su prikazani predstavljaju broj umrlih i živorođenih u 2015. godini u Srbiji, podaci Zavoda za statistiku, <http://www.stat.gov.rs/>.



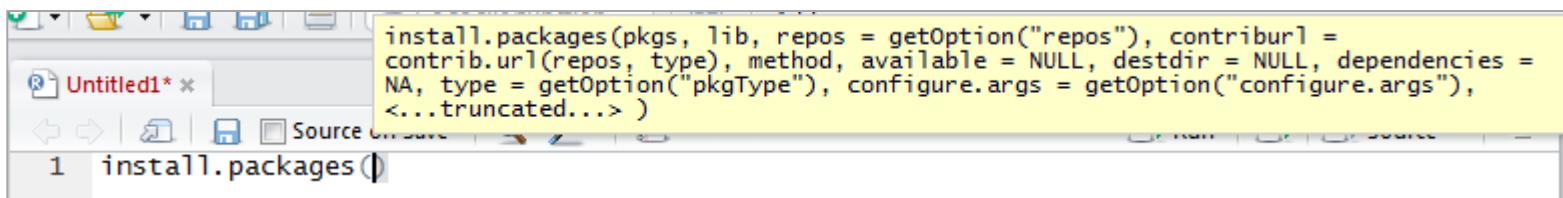
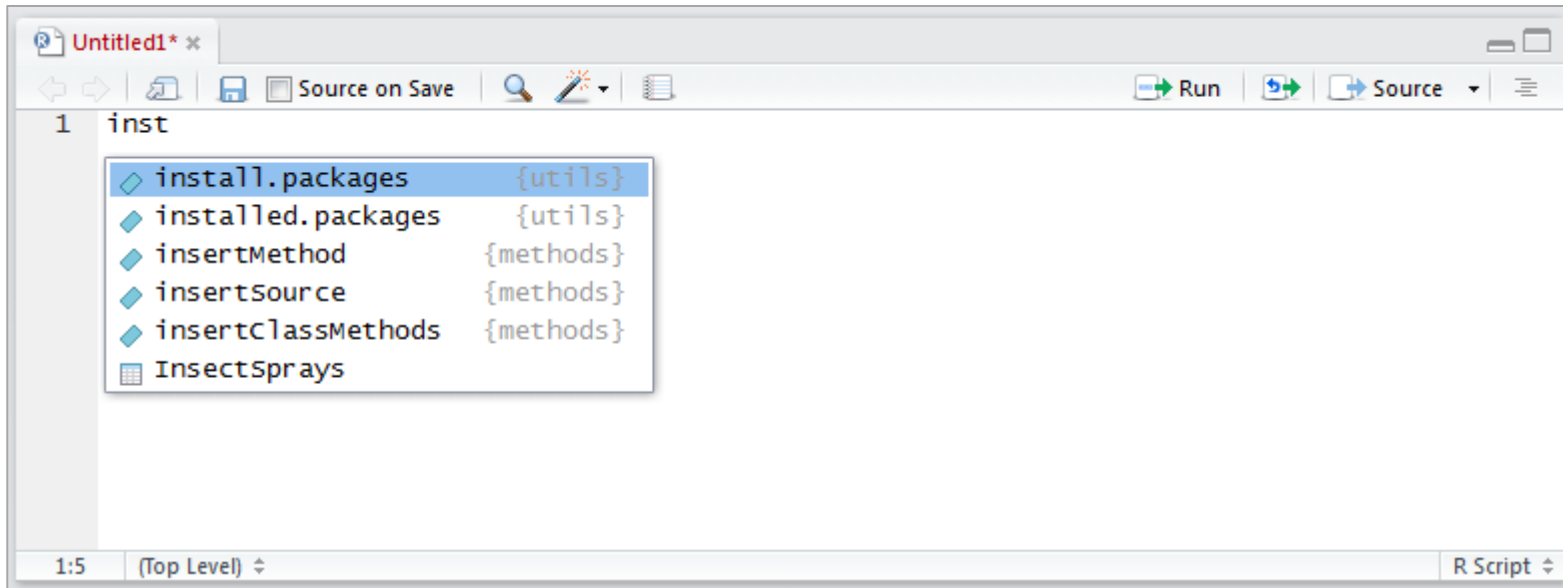
R Studio

- U *Environment* kartici je moguće pregledati kreirane objekte u R-u.
- U *Plot* prozoru se prikazuju grafici i iz padajućeg menija moguće je odabrati *Export* funkciju i snimiti sliku u datoteku (.pdf ili neka druga ekstenzija).
- Skript je moguće osim sa ekstenzijom .R snimiti i sa .Rmd (eng. *R markdown*) ekstenzijom. O tome kasnije ...
- Moguće je pokretati kod deo po deo odabirom odgovarajućih opcija iz padajućeg menije *Code*.

R paketi

- R studio je IDE (eng. *Integrated Development Environment*) za programski jezik R.
- Programski jezik R se može koristiti i bez R Studio-a, ali je bolje sa njim ... na TOBS predmetu se koristi R Studio.
- R dolazi sa relativno malim brojem osnovnih paketa (eng. *R packages*), pa je druge pakete potrebno instalirati.
- Komanda koja se može koristiti za instalaciju paketa je: *install.packages*.
- Ukucati u konzoli R Studio-a *install.packages*. Primetiti da je R Studio “pametno” okruženje -> *auto complete* opcija.
- Kada se ukuca ime paketa, onda ova komanda omogućava R-u da se poveže sa CRAN-om, da nađe paket i da ga instalira na računaru.
- Da bi se funkcije paketa mogle koristiti, potrebno je iskoristiti funkciju *library()* i (bez znaka navoda) uneti ime paketa koji se koristi.
- DODATNO: Kako bi se iz .R skripte pokrenula jedna linija koda ili se ona označi, pa se klikne na dugme u gornjem desnom uglu *Editor*-a ili sa tastature *CTRL-ENTER* (<- vrlo korisna skraćenica)!

Instalacija paketa u R-u



- Komanda za instalaciju paketa:
 - primetiti da postoji *auto-complete* u R-u koji omogućava da se pozove odgovarajuća funkcija (ovo se odnosi i na argumente funkcija i na objekte koje kreira korisnik) i
 - ova funkcija se konektuje preko interneta, preuzima i instalira odgovarajući paket.
- Koristan link oko instalacija je: <https://github.com/genomicsclass/windows> (pristupljeno 09.03.2024).

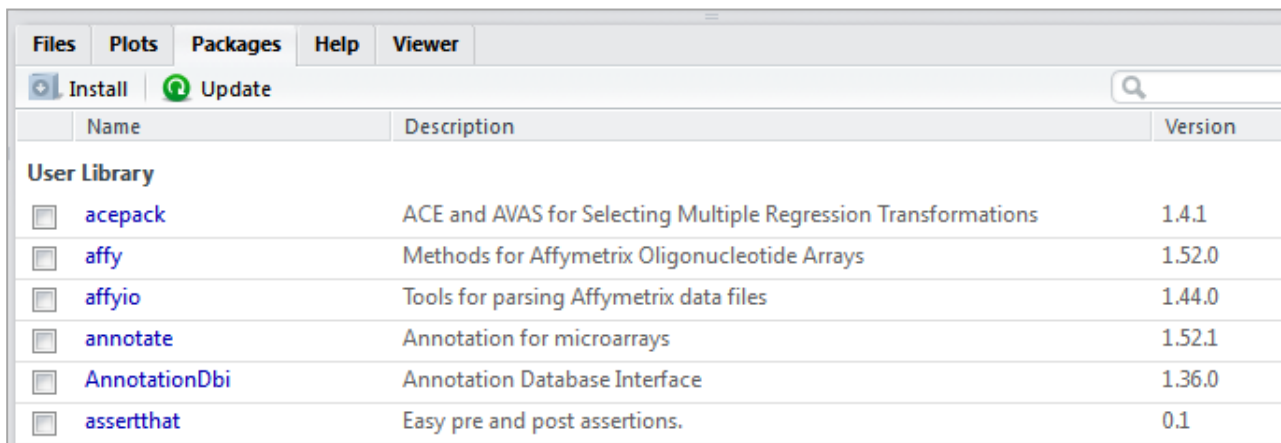
Drugi načini za instalaciju paketa

- Osim na CRAN-u, koji sadrži veliku biblioteku R paketa, moguće je naći pakete i na drugim sajtovima (npr. <https://www.bioconductor.org/>, pristupljeno 09.03.2024).
- Na kraju časa će se instalirati paket SWIRL koji će se koristiti za interaktivno učenje R jezika (<https://swirlstats.com/>, pristupljeno 09.03.2024).



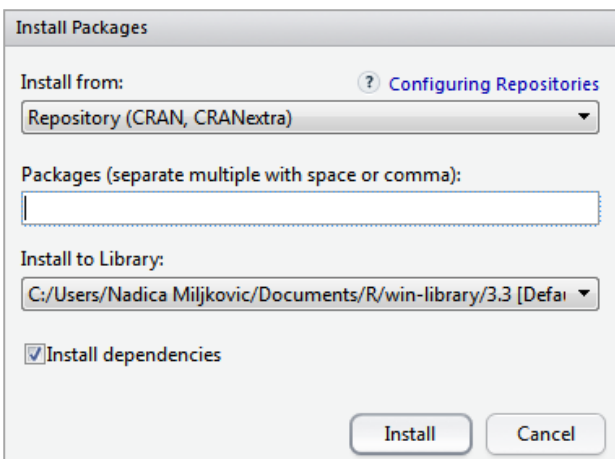
{swirl}

Drugi načini za instalaciju paketa



The screenshot shows the 'Packages' tab in R Studio. It features a search bar and two buttons: 'Install' and 'Update'. Below is a table with columns for 'Name', 'Description', and 'Version'. The table lists several packages under the 'User Library' section.

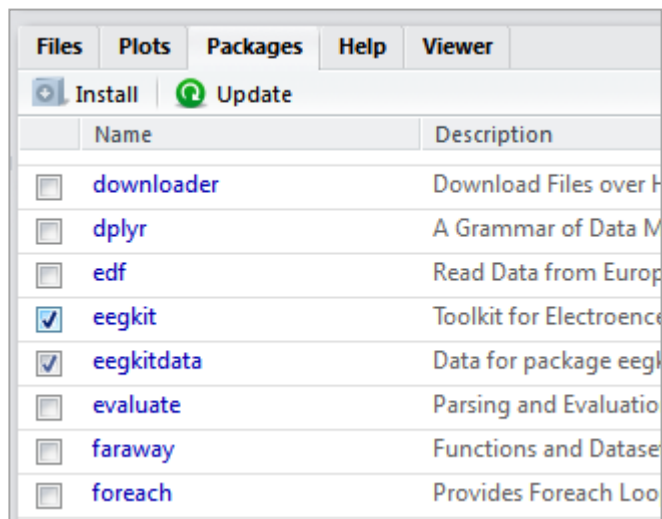
Name	Description	Version
<input type="checkbox"/> acepack	ACE and AVAS for Selecting Multiple Regression Transformations	1.4.1
<input type="checkbox"/> affy	Methods for Affymetrix Oligonucleotide Arrays	1.52.0
<input type="checkbox"/> affyio	Tools for parsing Affymetrix data files	1.44.0
<input type="checkbox"/> annotate	Annotation for microarrays	1.52.1
<input type="checkbox"/> AnnotationDbi	Annotation Database Interface	1.36.0
<input type="checkbox"/> assertthat	Easy pre and post assertions.	0.1



The 'Install Packages' dialog box is shown. It includes a dropdown menu for 'Install from:' set to 'Repository (CRAN, CRANextra)', a text input field for 'Packages (separate multiple with space or comma):', a dropdown menu for 'Install to Library:' set to 'C:/Users/Nadica Miljkovic/Documents/R/win-library/3.3 [Default]', and a checked checkbox for 'Install dependencies'. 'Install' and 'Cancel' buttons are at the bottom.

- U kartici *Packages* u interfejsu R Studio-a moguće je pregledati sve trenutno instalirane pakete na računaru kao na slici na gornjem panelu.
- Klikom na ikonicu *Install*, otvara se prozor kao na slici, gde je moguće izabrati i instalirati željeni paket kao na slici na donjem panelu.

Učitavanje paketa

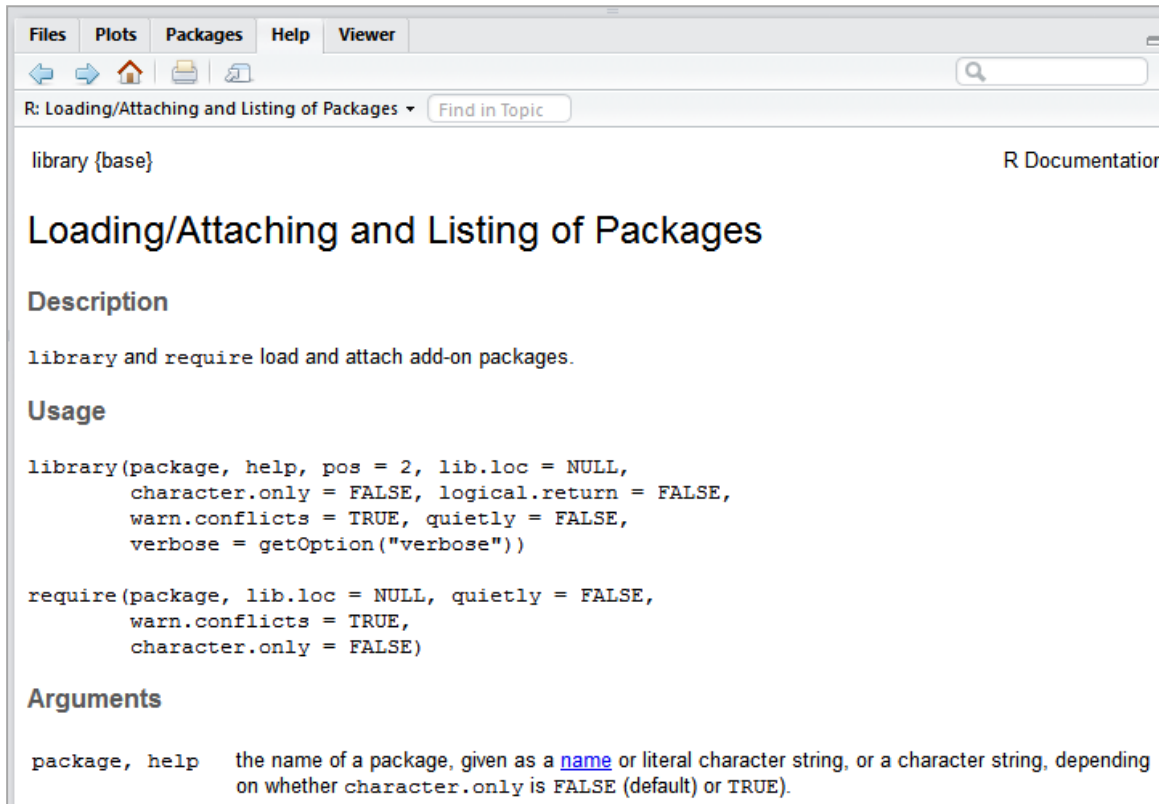


	Name	Description
<input type="checkbox"/>	downloader	Download Files over H
<input type="checkbox"/>	dplyr	A Grammar of Data M
<input type="checkbox"/>	edf	Read Data from Europ
<input checked="" type="checkbox"/>	eegkit	Toolkit for Electroence
<input checked="" type="checkbox"/>	eegkitdata	Data for package eegl
<input type="checkbox"/>	evaluate	Parsing and Evaluatio
<input type="checkbox"/>	faraway	Functions and Datas
<input type="checkbox"/>	foreach	Provides Foreach Loo

```
> library("eegkit", lib.loc="~/R/win-library/3.3")  
Loading required package: bigsplines  
Loading required package: eegkitdata  
Loading required package: ica  
Loading required package: rgl
```

- Ako je potrebno koristiti funkcionalnost nekog paketa, onda je potrebno uključiti te funkcije paketa pozivom komande *library()*.
- Za “isključivanje” funkcija nekog paketa koristiti komandu *detach()*. Verovatno Vam nikada neće ni biti potrebno.
- Ovo je sve moguće postići u kartici *Packages* interfejsa u R Studiju tako što se odgovarajući paket naznači (štiklira) ili ne. Pogledati sliku i obratiti pažnju na kod koji se automatski generiše kao posledica uključenja/isključenja odgovarajućih paketa.
- Poželjno je koristiti *library()* opciju zbog računarske reproducibilnosti.

Help (?) – najjednostavniji način



The screenshot shows the R Help window for the `library` function. The window title is "R: Loading/Attaching and Listing of Packages". The main content area displays the following information:

```
library {base} R Documentation
```

Loading/Attaching and Listing of Packages

Description

`library` and `require` load and attach add-on packages.

Usage

```
library(package, help, pos = 2, lib.loc = NULL,  
        character.only = FALSE, logical.return = FALSE,  
        warn.conflicts = TRUE, quietly = FALSE,  
        verbose = getOption("verbose"))
```

```
require(package, lib.loc = NULL, quietly = FALSE,  
        warn.conflicts = TRUE,  
        character.only = FALSE)
```

Arguments

`package`, `help` the name of a package, given as a [name](#) or literal character string, or a character string, depending on whether `character.only` is `FALSE` (default) or `TRUE`.

```
> ?library  
> |
```

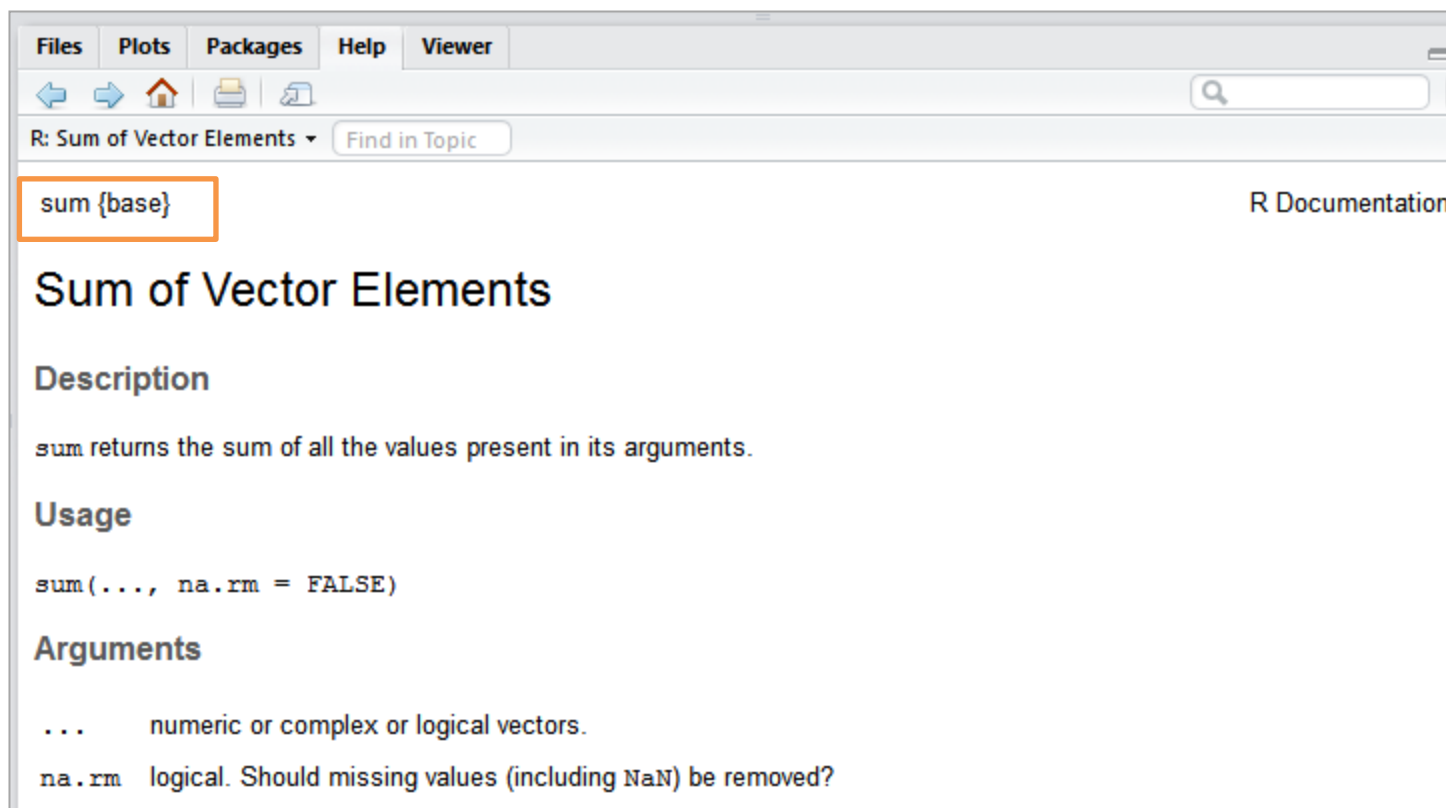
- Za uputstvo kako se koristi odgovarajuća komanda/funkcija pokrenuti liniju koda sa `?library`.
- U kartici *Help* na interfejsu će se pojaviti uputstvo za korišćenje odgovarajuće komande, odnosno R dokumentacija.

Help – drugi načini

```
>  
> help.start()  
If nothing happens, you should open  
'http://127.0.0.1:16941/doc/html/index.html' yourself  
> |
```

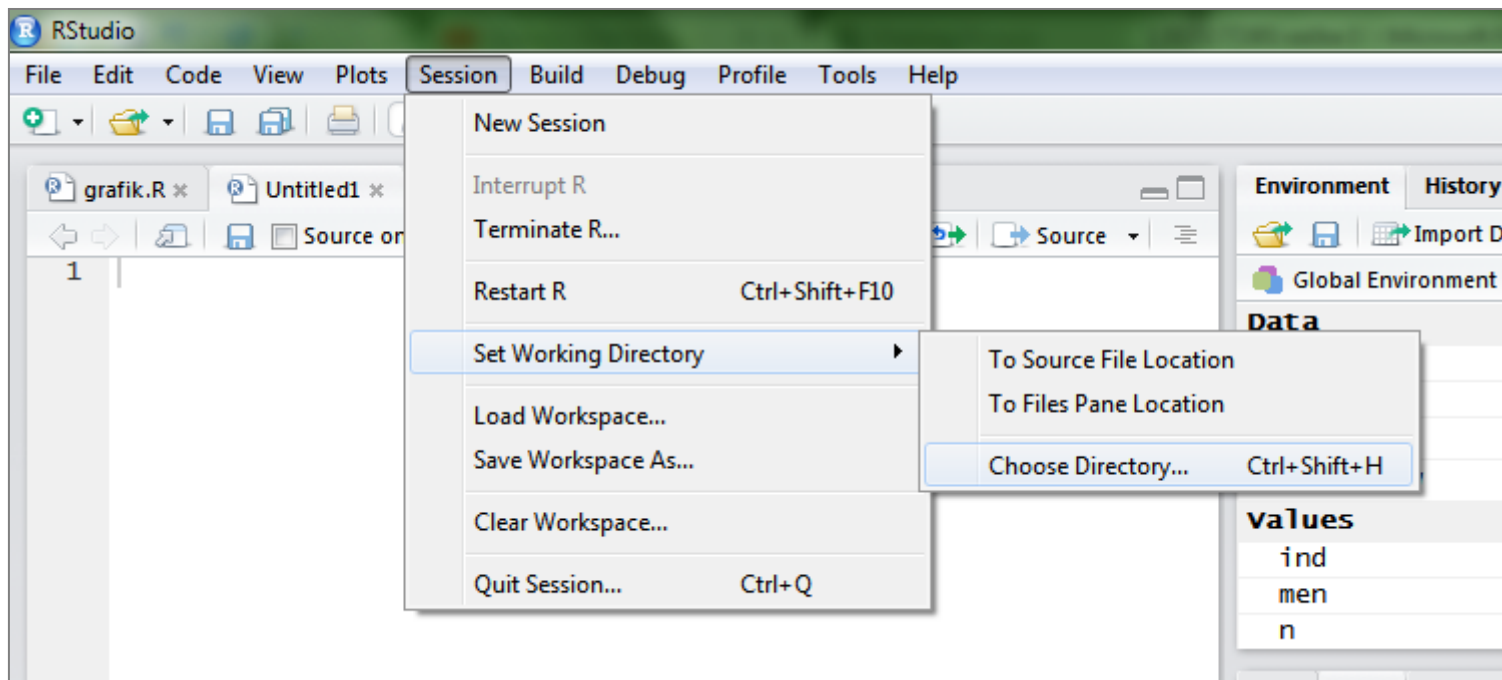
- Moguće je u konzoli osim komande *?sum*, ukucati i *help(sum)*. Efekat je isti, otvoriće se prozor iz R dokumentacije koji se odnosi na sumiranje elemenata niza.
- Moguće je iz komandnog prozora pokrenuti i pretragu po ključnim rečima. Na primer, za ključnu reč *regression*, komanda je *help.search("regression")*.
- Za HTML verziju R dokumentacije moguće je ukucati komandu *help.start()*.
- Ovim opcijama je moguće pristupiti i u *Help* kartici R Studio interfejsa.
- Moguće je koristiti opcije: *Next topic*, *Previous topic*, *Show R help*, *Print topic*, *Show in new window*. Takođe, postoje i dve opcije za pretragu: pretraga po temi i pretraga celokupne dokumentacije.

Help kartica



- Primetiti da je otvorena opcija za dokumentaciju o `sum()` funkciji.
- Na slici je uokviren naziv paketa u kome se nalazi funkcija za koju je izlistana dokumentacija.

R Studio – radni direktorijum



- Prvo je potrebno proveriti kako je podešen radni direktorijum (*Working directory*) -> ovo je važno za manipulaciju podacima koji su upisani u datoteke i isto tako za memorijski prostor koji je određen za čuvanje promenljivih.
- Da bi se proverio trenutni direktorijum potrebno je ukucati `getwd()` u komandni prozor R-a.
- Ako bi se pokušalo sa otvaranjem nekog dokumenta koji se ne nalazi u radnom direktorijumu, R bi javio grešku (da se ta datoteka ne nalazi u radnom direktorijumu).
- Na slici je prikazano kako je moguće promeniti radni direktorijum. Umesto opcije sa slike moguće je koristiti i funkciju `setwd()`.
- Korisna linija koda: `setwd(dirname(rstudioapi::getSourceEditorContext())$path)` – potrebno je instalirati `rstudioapi` biblioteku (pogledati komentare u rešenju vežbe 1).

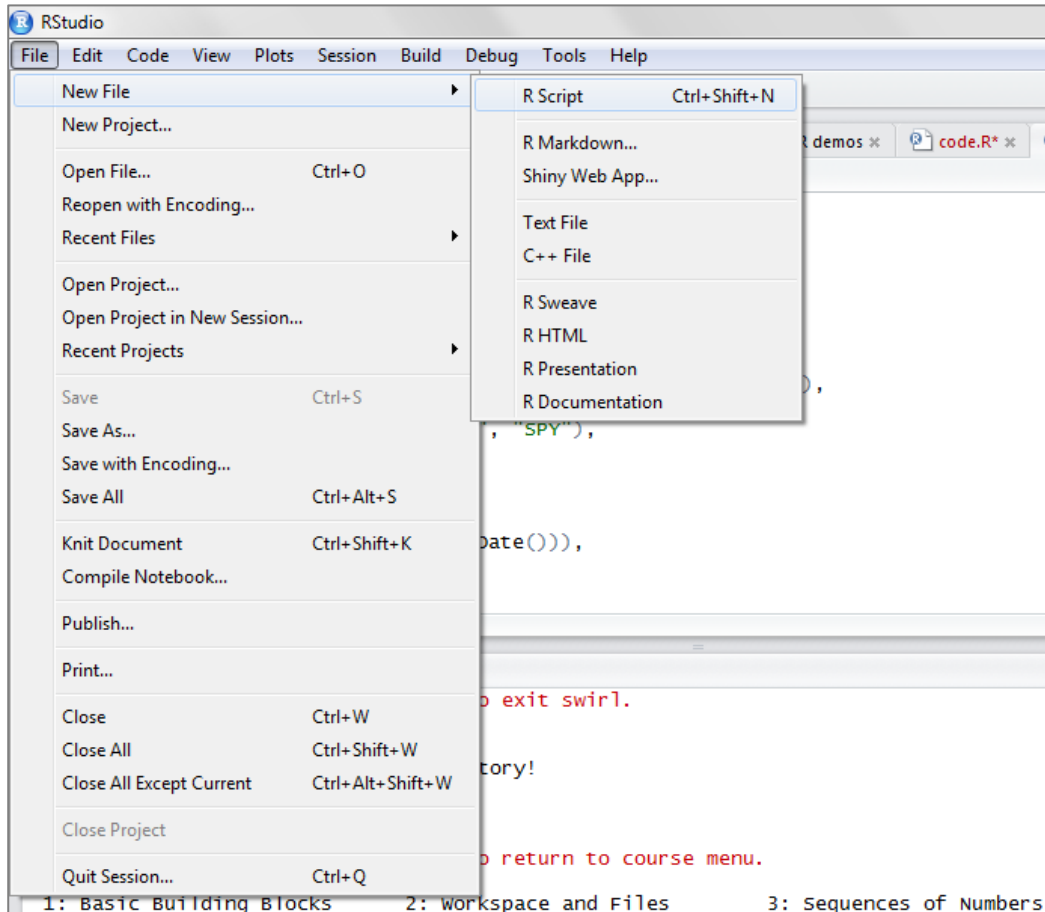
R Studio – skripta

- Kucanjem komande *dir()* dobija se lista svih datoteka koji se nalaze u radnom direktorijumu.
- Potrebno je sve komande pisati u tekstualnom editoru tj. skripti (što je jednostavnije nego u komandnom prozoru i važno zbog računarske reproducibilnosti).

Modifikovana slika: Snow od Richard Gillin; Flickr [link](#); CC BY-SA 2.0



R Studio – otvaranje tekstualnog editora



- Na slici levo je pokazano kako se otvara tekstualni editor (*R Script*).
- Skript/datoteka se snima sa ekstenzijom “.R”.

Unos podataka

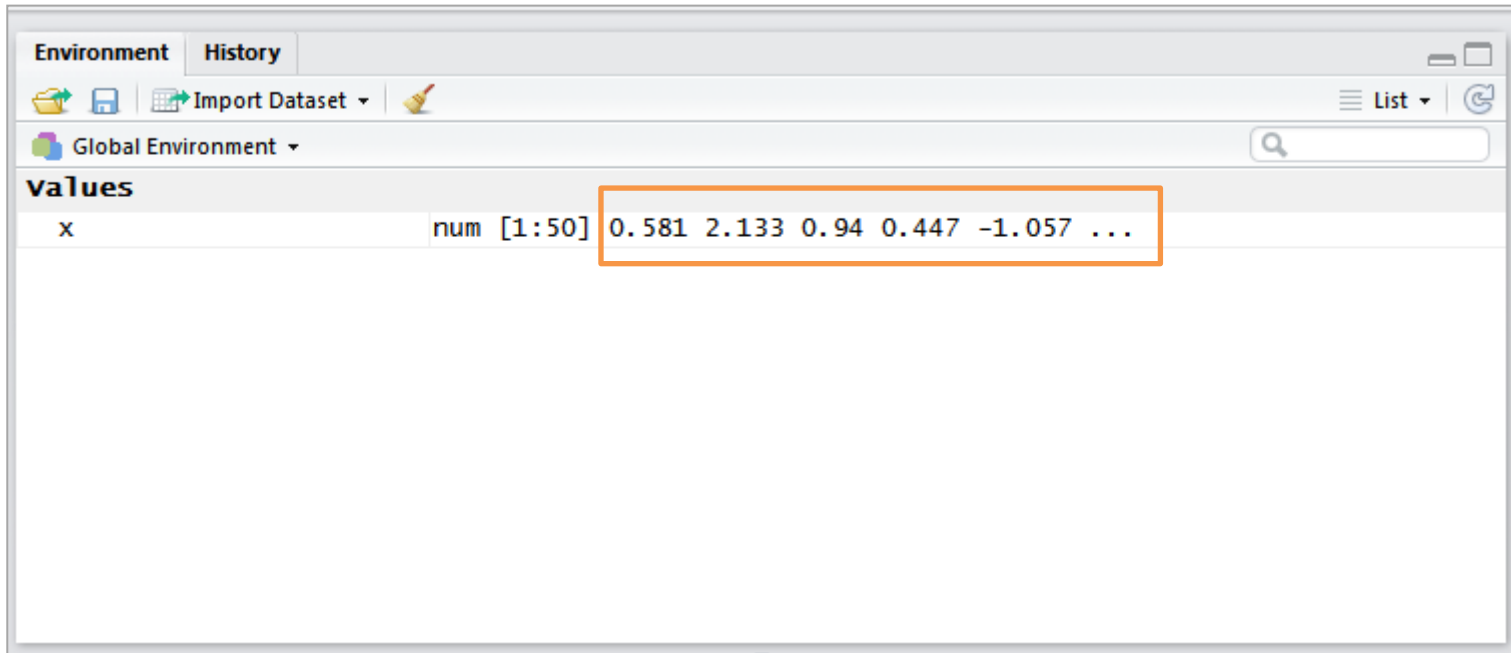
```
> x <- 5
> x
[1] 5
> print(x)
[1] 5
> x <- 1:10 # x je promenljiva koja ima vrednosti 1, ..., 10
> x
[1] 1 2 3 4 5 6 7 8 9 10
> print(x)
[1] 1 2 3 4 5 6 7 8 9 10
> |
```

- U komandnom prozoru (konzoli) je moguće pisati izraze kao u primeru sa slike i prikazivati vrednosti odgovarajućih promenljivih.
- R proverava da li je izraz kompletan prilikom kucanja.
- Oznaka taraba “#” označava da je sve sa desne strane ovog znaka komentar (uključujući i “#”).
- **Primetiti da je oznaka/operator za dodeljivanje vrednosti u R-u jednaka “<-”. Po želji, umesto znaka za dodelu vrednosti, u R-u se može koristiti i znak jednakosti “=”.**
- Postoji *autoprint*, ali i *explicit print*. Implicitno i eksplicitno štampanje.

Automatsko/implicitno i eksplicitno štampanje

- *Automatic print* – ako se upiše ime promenljive u kojoj su smešteni određeni podaci (i pokrene ta linija koda), onda se realizuje automatsko štampanje (prikaz na ekranu tj. konzoli).
- *Explicit print* – ako se koristi funkcija *print()* koja kao ulazni parametar može imati promenljivu koja se eksplicitno štampa tj. ispisuje na konzoli.
- *Automatic print* je moguće i postići pozivom funkcije bez dodeljivanja vrednosti (izlaza funkcije) nekoj promenljivoj.
- **ZADATAK: Šta radi funkcija *rnorm()*? Generisati 5 (pseudo)slučajnih brojeva pozivom ove funkcije *rnorm(5)* i dodeliti rezultat promenljivoj *x*. Primeniti automatsko i eksplicitno štampanje vrednosti promenljive na ekranu (konzoli).**

Zadatak – dodatno



- Primetiti da se unutar kartice *Environment* (okruženje) pojavljuje i promenljiva *x*.
- Pokušajte da odgovorite na pitanja:
 - Kog je tipa promenljiva *x*?
 - Koliko elemenata ima promenljiva *x*? Da li je to očekivano?
 - Šta označavaju uokvireni brojevi na slici?

Ispis podataka na ekranu

Šta znače ove oznake?

```
> y <- 1.50
> print(y)
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
[34] 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
> y
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33
[34] 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
> |
```

Šta će se ispisati na ekranu, ako se prethodni zadatak ponovi za 50 generisanih izlaza funkcije *rnorm()*? Obratiti pažnju da prikaz sa slike ne odgovara zadatku!

Šta se menja u kartici *Environment*?





Ako se zadatak ponovi i treći put, za isti naziv promenljive, ali za tri generisana izlaza funkcije, koja je dimenzija x ? (*HINT*: pogledati *Environment*)

Operacije i osnovni tipovi podataka

```
> c(0.5, 1.2, 0.8, 0.9)
[1] 0.5 1.2 0.8 0.9
> x <- c(0.5, 1.2, 0.8, 0.9)
> x
[1] 0.5 1.2 0.8 0.9
> sum(x)
[1] 3.4
> 0.5 + 1.2 + 0.8 + 0.9
[1] 3.4
> |
```

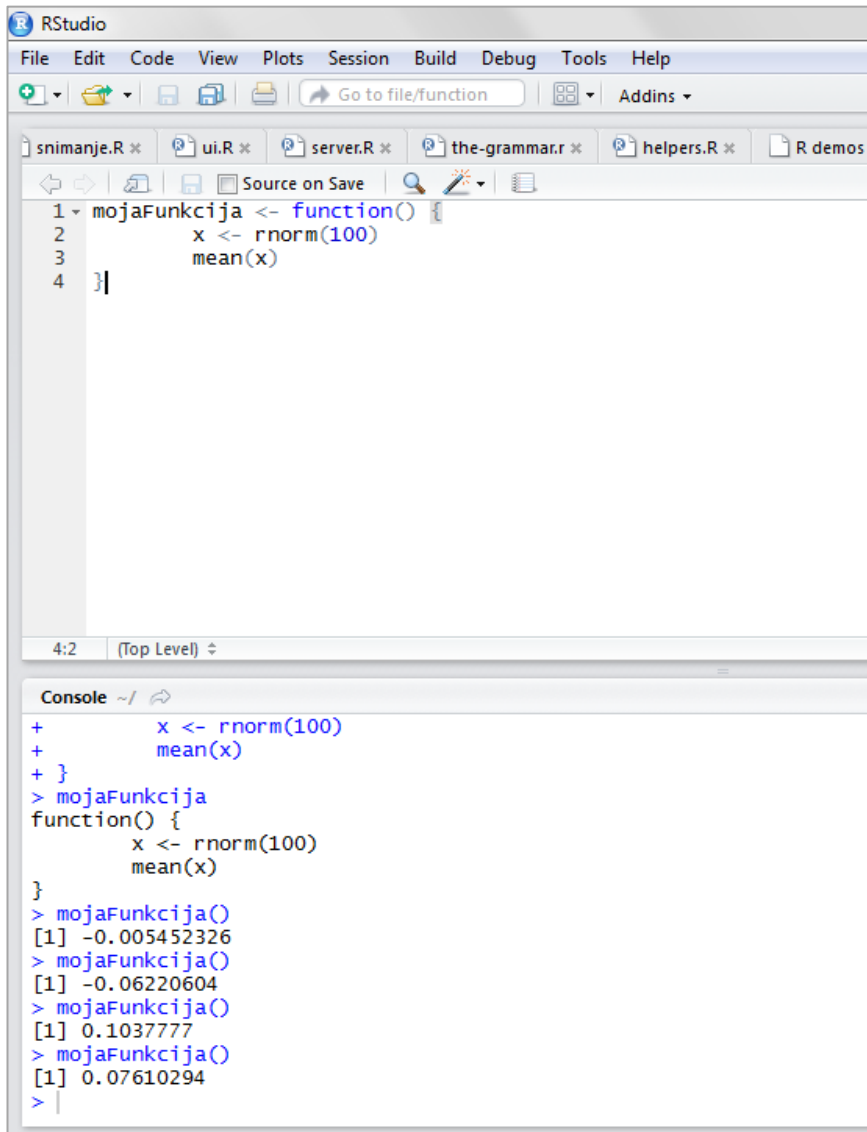
- Na sledećem času vežbi se rade osnovne operacije, kao i osnovni tipovi podataka ...
- Međutim, za prvi čas, potrebno je da se napomene da se niz brojeva od npr. brojeva 0.5, 1.2, 0.8 i 0.9 formira pomoću funkcije `c()` kao na slici.

Sabiranje nizova

```
Console ~/      
>  
>  
>  
> niz1 <- c(1, 1, 1, 1, 1, 1)  
> niz2 <- c(2, 3, 4)  
> niz1 + niz2  
[1] 3 4 5 3 4 5  
> niz3 <- c(2, 3, 4, 5)  
> niz1 + niz3  
[1] 3 4 5 6 3 4  
warning message:  
In niz1 + niz3 :  
  longer object length is not a multiple of shorter object length  
>
```

- Sabiranje dva niza različitih dužina ima rezultat kao na slici.
- Obratiti pažnju na upozorenje kada dimenzija jednog niza nije umnožak drugog.

Pisanje funkcija



The screenshot shows the RStudio interface. The source editor contains the following R code:

```
1 mojaFunkcija <- function() {  
2   x <- rnorm(100)  
3   mean(x)  
4 }
```

The console shows the execution of the function definition and several calls to the function, resulting in different numerical outputs:

```
+ x <- rnorm(100)  
+ mean(x)  
+ }  
> mojaFunkcija  
function() {  
  x <- rnorm(100)  
  mean(x)  
}  
> mojaFunkcija()  
[1] -0.005452326  
> mojaFunkcija()  
[1] -0.06220604  
> mojaFunkcija()  
[1] 0.1037777  
> mojaFunkcija()  
[1] 0.07610294  
> |
```

- **ZADATAK:** Napisati jednostavnu funkciju.
- Cilj funkcije je da generiše 100 normalno raspodeljenih pseudoslučajnih brojeva i dodatno:
 - funkcija nema ulaze i
 - funkcija na izlazu “vraća” srednju vrednost tih brojeva.
- Pre korišćenja funkcije, potrebno je učitati funkciju u R.
- Funkcija se pokreće/učitava (liniju po liniju) sa komandama sa tastature CTRL+ENTER.
- Pokrenuti više puta ovu funkciju u konzoli.
- Zašto se rezultat razlikuje?

Pisanje i pokretanje funkcija

- Da bi se pokrenula funkcija potrebno ju je učitati korišćenjem komande *load()*.
- *ls()* funkcija pokazuje da li se nešto nalazi u *Environment*-u.
- Nakon toga moguće je pokrenuti funkciju i dobiti rezultat kao na prethodnom slajdu.
- Može se iskopirati i u komandnom prozoru/konzoli, pa tek onda pokrenuti.
- Proveriti sa *ls()* funkcijom kada se učita funkcija da li se nalazi u radnom okruženju (*Workspace*-u), pre korišćenja.
- Ako se snima tekstulani editor sa funkcijom ekstenzija je “.R”.
- Ime funkcije ne mora da odgovara imenu tekstulanog editora u kome se funkcija nalazi, ali je nekada zgodnije da imena budu ista.
 - Ovo je razlika u odnosu na Matlab/GNU Octave.
- Može se dodati u *Workspace* sa komandom *source(“mojaFunkcija.R”)*.

Zadatak #2

- ZADATAK: Dodati drugu funkciju *mojaFunkcija2()*, pa onda sa *ls()* proveriti šta se sve nalazi u *Environment*-u i onda testirati obe funkcije. Neka ta druga funkcija dodaje niz od 5 brojeva koji su normalno raspodeljeni (šum) nekom broju (DC komponenta) koji predstavlja ulazni argument funkcije. Pretpostaviti da ulazni parametar može da bude bilo kojih dimenzija i da mu je podrazumevana vrednost vektor/niz dužine 1 koji ima vrednost 0.
- Zašto funkciju vraća kao rezultat različite vrednosti?
- Kako funkcija “zna” koji rezultat je potrebno vratiti kao konačan?

Funkcije i komande za danas

- `install.packages()`
- `library()`
- `detach()`
- `?`
- `help()`
- `help.search()`
- `help.start()`
- `sum()`
- `getwd()`
- `setwd()`
- `dir()`
- `print()`
- `<-`, `-`, `+`
- `#`
- `rnorm()`
- `c()`
- `mean()`
- `function()`
- `{}`
- `ls()`

Do sledećeg časa ...

- Nijedan jezik se ne uči čitanjem rečnika, ali rečnici mogu biti korisni prilikom učenja. Primeri su:
 - <https://cran.r-project.org/doc/contrib/Short-refcard.pdf> (za sada, pristupljeno 09.03.2024.)
 - <https://posit.co/resources/cheatsheets/> (za kasnije, pristupljeno 09.03.2024.)
- Za učenje jezika je najbolje vežbanje.
- Sada, Pokrenuti SWIRL i uraditi prvu lekciju pod nazivom *Basic Building Blocks*. Na kraju lekcije, nije potrebno slati svoje rezultate dalje.

Interaktivno učenje



```
Console ~/  
| course installed successfully!

| Please choose a course, or type 0 to exit swirl.

1: R Programming
2: Take me to the swirl course repository!

selection: 1

| Please choose a lesson, or type 0 to return to course menu.

1: Basic Building Blocks      2: Workspace and Files      3: Sequences of Numbers
4: Vectors                    5: Missing Values           6: Subsetting Vectors
7: Matrices and Data Frames  8: Logic                     9: Functions
10: lapply and sapply        11: vapply and tapply       12: Looking at Data
13: Simulation                14: Dates and Times         15: Base Graphics

selection:
```

- Instalirati swirl paket (<http://swirlstats.com/>, pristupljeno 09.03.2024.) koji služi za interaktivno učenje programskog jezika R.
- Učitati ovaj paket i po potrebi instalirati kurs `install_course("R Programming E")`.
- Pratiti uputstva i pokrenuti interaktivno vežbanje.
- Uraditi lekciju iz vežbanja 1 (R Programming E: *Basic Building Blocks*).