European Social Survey

R programiranje u društvenim naukama

Radmila Veličković,

22.2.2019, Beograd

O meni

Master Metodologija anketiranja i javno mnjenje, Univerzitet u Nojšatelu (major)

Master Data Science, Univerzitet u Cirihu (minor)

Master in Statistics

ESRA (European Survey Research Association)

RLadies Belgrade/ RLadies Lausanne/ Swiss-AI/WiMLDS

BernR

Zürich Women in Machine Learning and Data Science

A vi ste?

Vaše iskustvo sa  -om?

Zašto  ?

Struktura kursa		
Prvi deo	**Drugi deo**	**Treći deo**
1.Uvod u R programski jezik 2. R okruženje 3. Šta sve R može? 4. Prvi koraci u R-u 5. Importovanje file-ova (SPSS, STATA, csv, itd) 6. Prve analize podataka 7. Manipulacija podacima 8. Sređivanje seta podataka (nedostajuće vrednosti, transformacija varijablil i sl.)	9. Deskriptivna statistika 10. Statistički testovi 11. Prosta linearna regresija 12. Višestruka linearna regresija 13. Vizuelizacija podataka (Uvod u ggplot2) 14. Analiza korespondencije	15. R i ostali programi 16. Popularni paketi 17. Literatura za (samostalno) učenje 18. Šta dalje?

Da li biste nešto dodali/izbacili?

Sadržaj za danas

1. Uvod u R programski jezik
2. Šta sve R može?
3. Prvi koraci u R-u
4. Importovanje podataka
5. Manipulacija podacima (dplyr)

Uvod u R programski jezik

R kao programski jezik (1993.)

Preteča R-a je programski jezik S (Bell Laboratories)

Kreatori su Ross Ihaka i Robert Gentlemen

Objektno-orijentisano programiranje

Prvenstveno je bio namenjen statističarima

Danas je u širokoj upotrebi (analitičari tržišta, programeri, biolozi, hemičari, lingvisti, sociolozi, psiholozi, data inženjeri…)

Base R versus paketi (dodaci)

Uvod u R programski jezik

Popularan stiče jer je: open-source/jednostavan za upotrebu/ima ogroman community/fleksibilan/ima veliki izbor paketa

Vodi ga CRAN Community **(comprehensive R Archive Network)**

Ogroman doprinos naučne zajednice razvijanju R-a (razvoj paketa/prijava grešaka, rešavanja problema sa kojima se (početnici) suočavaju…)

stackoverflow

CRAN
Mirrors
What's new?
Task Views
Search

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
- Download R for (Mac) OS X
- Download R for Windows

R u društvenim naukama

Program koji se koristi u akademskim krugovima za analizu podataka

Sadrži pakete namenjene:

- psiholozima (psych, psychometric…)

- sociolozima (essurvey, AnthroTools, WeightedCluster, igraph…)

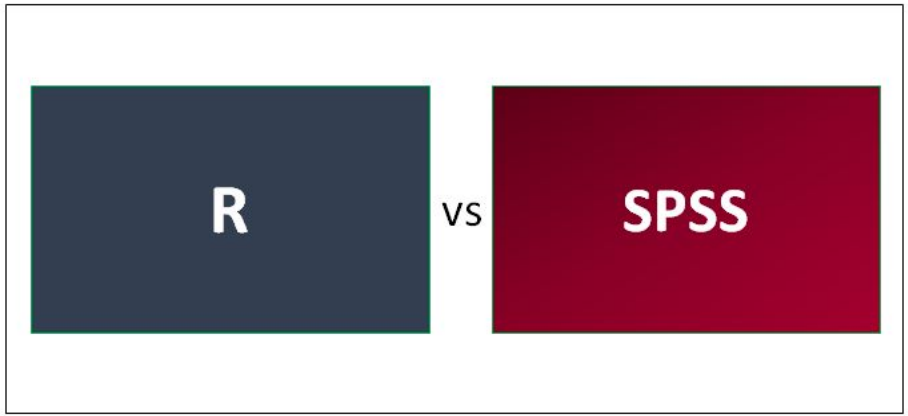
-lingivistima (languageR, corpus, twitteR…)

- i mnogim drugima

- omogućava reprodukovanje istraživanja (reproducible research)

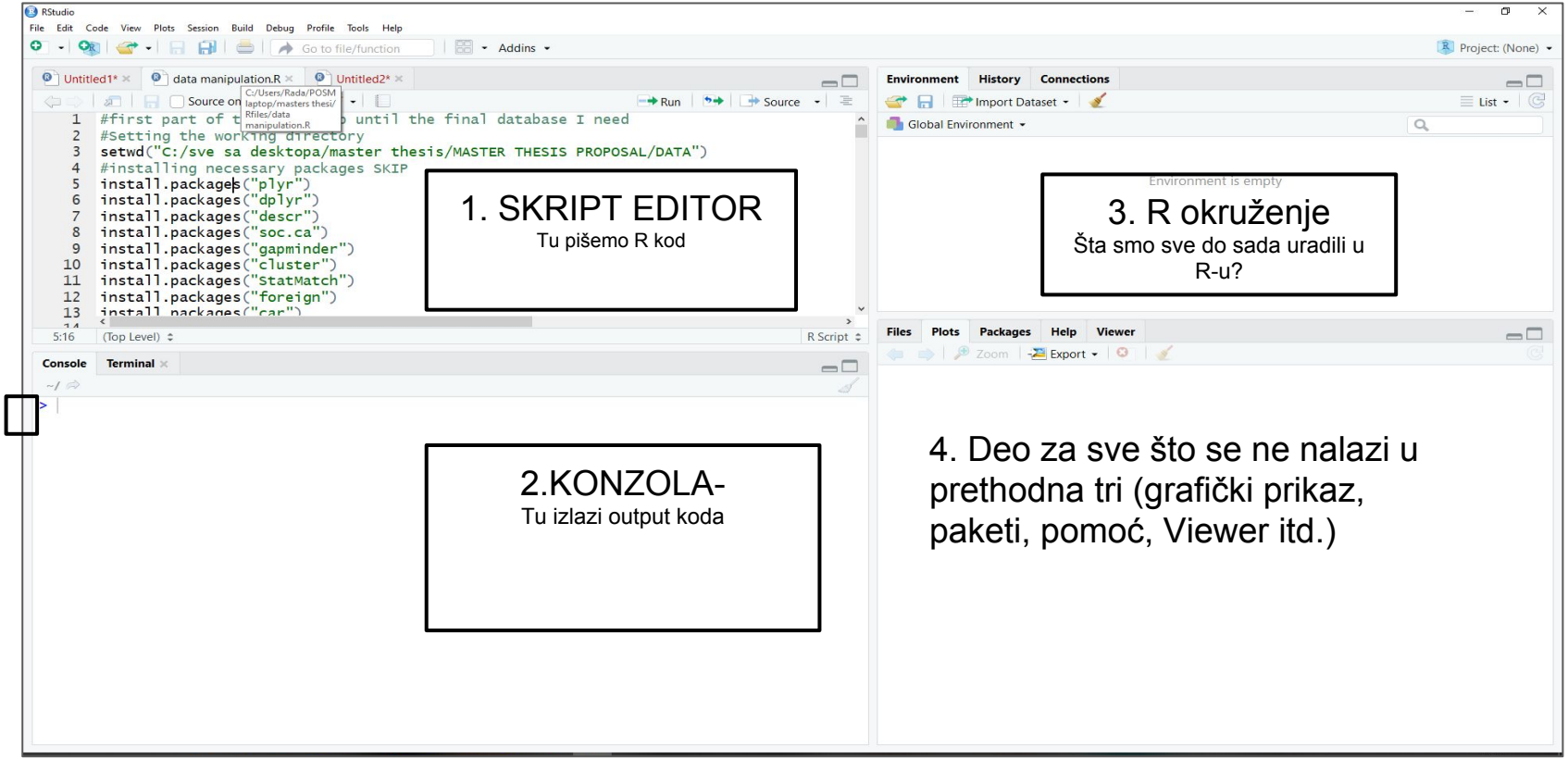
Uvod u R programski jezik

R vs R Studio


R
vs
SPSS

Šta sve R može?

Prvi koraci u R-u: Kako izgleda R Studio?



Prvi koraci u R-u: Osnovni pojmovi

U R-u kodiramo tj. pišemo **funkcije**

Funkcije su deo base R-a ili se nalaze u nekom paketu

Možemo ih i samostalno napisati (user-defined functions)

Primer: mean()

U zagradi funkcije uvek stoje **argumenti** funkcije.

Primer: mean (ESS$polintr)

Prvi koraci u R-u: Osnovni pojmovi

Funkcije možemo da smestimo u **objekat**

sv_polintr<-mean(ESS$polintr)

```
3  mean(ESS$polintr)
4
5  svpolinr<-mean(ESS$polintr)
6
```

<- assign

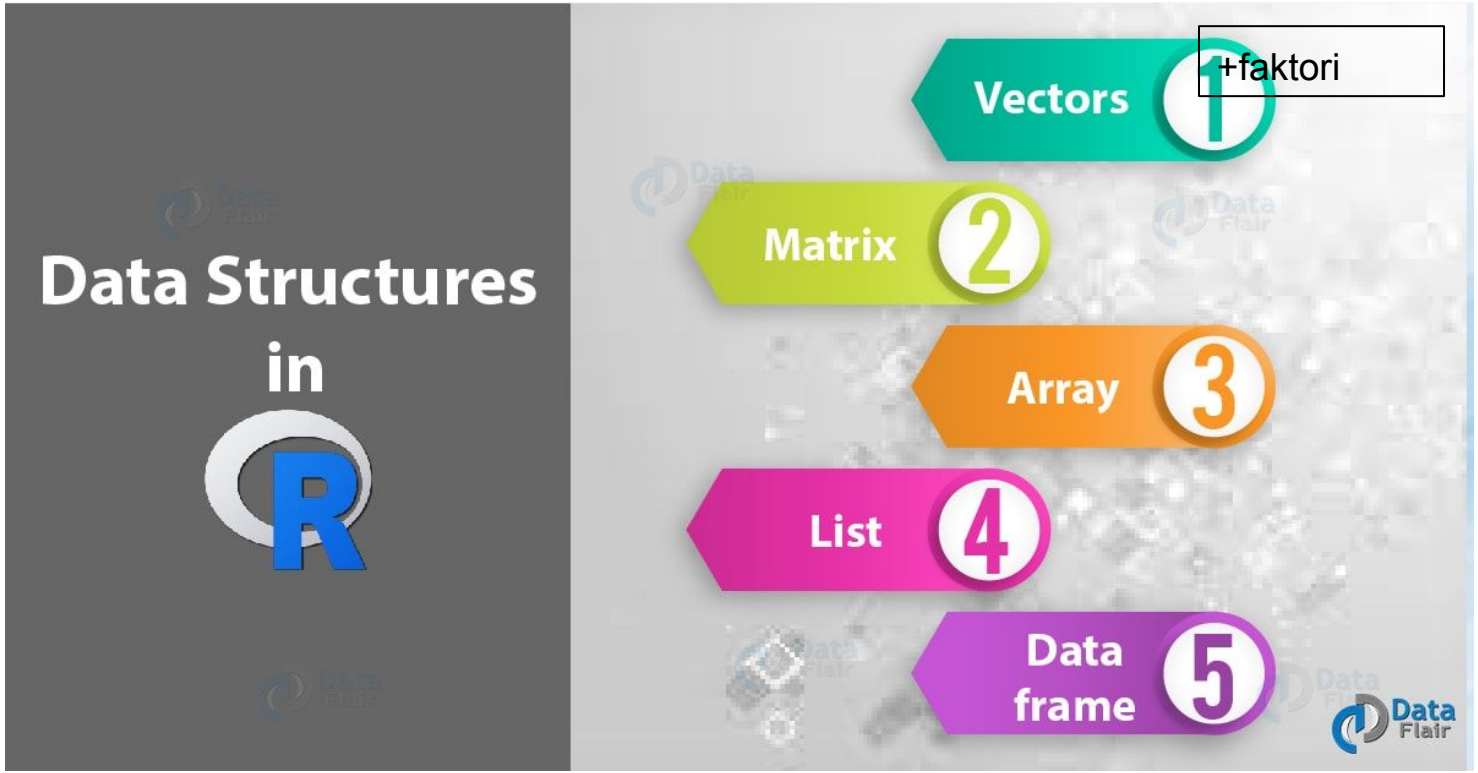
Kod/funkciju **izvršavamo** (to call a code/function) sa ctrl+Enter

služi za pisanje komentara

Da li ste svi instalirali R i RStudio?

https://cran.r-project.org/
https://www.rstudio.com/products/rstudio/download/

Strukture podataka



+faktori

Strukture i tipovi podataka: VEKTORI.

Nizovi elemenata koji omogućavaju unos podataka

Jedan vektor je jedan red ili kolona

Elementi se povezuju funkcijom **c()**

c() concatenate

```
moj_prvi_vektor<-c("Allow many to come and live here","Allow some", "Allow a few", "Allow none")
  moj_drugi_vektor<-c(400,140,130,50)
```

Vektori mogu biti: **atomski i u formi liste**

Strukture i tipovi podataka: MATRICE.

Matrice u matematičkom smislu

Podržava samo numeričke vrednosti

Sve operacije sa matricama (matrična algebra) moguća u base R-u

Funkcija matrix()

```
A=matrix(
  c(2, 4, 3, 1, 5, 7),
  nrow=3,
  ncol=2)
```

Strukture: FAKTORI.

NOMINALNE I ORDINALNE VARIJABLE (strings)

Broj postaje string kada dobije “3”

Razlikuje se od character u kontektu upotrebe ali značenje je isto

Moguće je transformisati jednu strukturu podatka u drugu

Strukture podataka: transformacija

is.atomic()

as.factor()

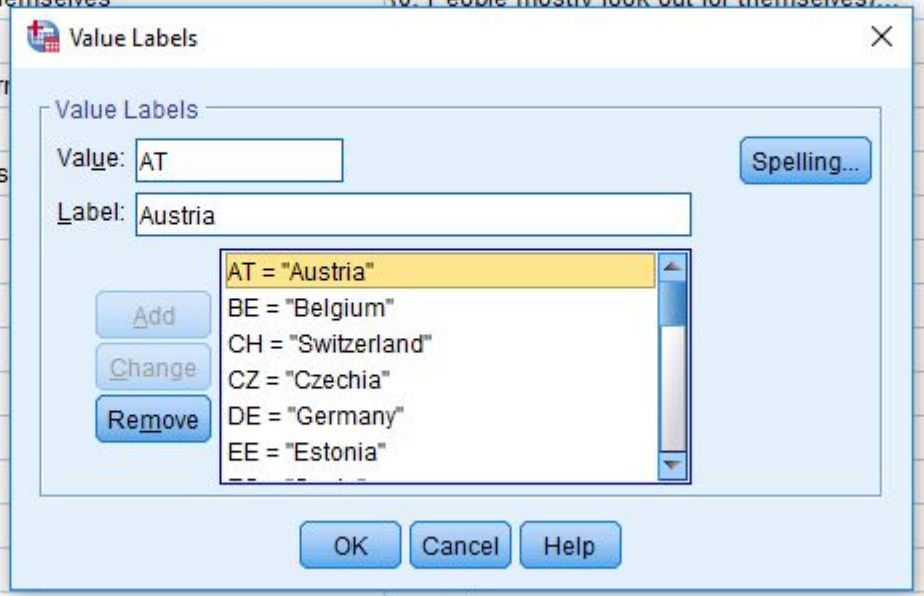
```
moj_prvi_vektor<-c("Allow many to come and live here","Allow some", "Allow a few", "Allow none")
is.atomic(moj_prvi_vektor)
moj_prvi_vektor<-as.factor(moj_prvi_vektor)
is.factor(moj_prvi_vektor)
```

Strukture podataka: data.frame

	name	essround	edition	proddate	idno	cntry	nwspol	netusoft	netustm	ppltrst	pplfair	pplhlp	polintr
1	ESS8e02_1	8	2.1	01.12.2018	1	AT	120	4	180	8	8	3	1
2	ESS8e02_1	8	2.1	01.12.2018	2	AT	120	5	120	6	6	5	1
3	ESS8e02_1	8	2.1	01.12.2018	4	AT	30	2	6666	5	6	4	3
4	ESS8e02_1	8	2.1	01.12.2018	6	AT	30	5	120	6	5	6	2
5	ESS8e02_1	8	2.1	01.12.2018	10	AT	30	5	180	5	5	7	3
6	ESS8e02_1	8	2.1	01.12.2018	11	AT	60	5	120	3	5	4	2
7	ESS8e02_1	8	2.1	01.12.2018	12	AT	15	2	6666	7	7	6	3
8	ESS8e02_1	8	2.1	01.12.2018	13	AT	45	4	30	7	7	7	3
9	ESS8e02_1	8	2.1	01.12.2018	14	AT	10	5	120	9	10	10	4
10	ESS8e02_1	8	2.1	01.12.2018	15	AT	60	4	120	5	3	4	2
11	ESS8e02_1	8	2.1	01.12.2018	16	AT	30	1	6666	2	0	0	3

Showing 1 to 12 of 44,387 entries

Internet use, how much time on typical day, in minutes	{6666, Not applicable}...	6666 - 9999
Most people can be trusted or you can't be too careful	{0, You can't be too careful}...	77, 88, 99
Most people try to take advantage of you, or try to be fair	{0, Most people try to take advantage of me}...	77, 88, 99
Most of the time people helpful or mostly looking out for themselves	{0, People mostly look out for themselves}...	77, 88, 99
How interested in politics		7, 8, 9
Political system allows people to have a say in what govern		7, 8, 9
Able to take active role in political group		7, 8, 9
Political system allows people to have influence on politics		7, 8, 9
Confident in own ability to participate in politics		7, 8, 9
Trust in country's parliament		77, 88, 99
Trust in the legal system		77, 88, 99
Trust in the police		77, 88, 99
Trust in politicians		77, 88, 99
Trust in political parties		77, 88, 99
Trust in the European Parliament		77, 88, 99
Trust in the United Nations		77, 88, 99
Voted last national election		7, 8, 9
Party voted for in last national election, Austria	{1, SPÖ}...	66 - 99
Party voted for in last national election, Belgium	{1, Groen!}...	66 - 99



ited or you can't be too careful	{0, You can't be too careful}...	77, 88, 99	9
advantage of you, or try to be fair	{0, Most people try to take advantage of me}...	77, 88, 99	9
helpful or mostly looking out for themselves	{0, People mostly look out for themselves}...	77, 88, 99	8
s		7, 8, 9	9
people to have a say in what goverr		7, 8, 9	10
n political group		7, 8, 9	10
people to have influence on politics		7, 8, 9	10
to participate in politics		7, 8, 9	10
nent		77, 88, 99	9
n		77, 88, 99	9
		77, 88, 99	9
		77, 88, 99	9
		77, 88, 99	9
arliament		77, 88, 99	8
ons		77, 88, 99	8
ion		7, 8, 9	6
tional election, Austria	{1, SPÖ}...	66 - 99	9
tional election, Belgium	{1, Groen!}...	66 - 99	9
tional election, Switzerland	{1, Swiss People's Party}...	66 - 99	9