

DODATNI KONCEPTI R

Hrvatski studiji

dr.sc. Luka Šikić

Preddiplomski studij sociologije

14 listopad, 2019

CILJEVI PREDAVANJA

- ▶ Što su paketi
- ▶ Radni prostor
- ▶ Učitavanje i spremanje podataka
- ▶ Specijalne vrijednosti, nazivi i klase varijabli
- ▶ Složeniji tipovi varijabli: faktori, podatkovni okviri, liste

RAD SA PAKETIMA

- ▶ paket(knjižnica) je skup funkcija(podataka i sl.) koje je netko napisao i ponudio na korištenje
- ▶ paketi su javno dostupni (Open Source)
- ▶ paket je potrebno jednom instalirati, a učitati se mora pri svakom otvaranju sesije - u okviru kolegija ćemo uglavnom koristiti nekoliko paketa: *base R*, *lsr*, *ggplot2*, *psych*, *car* i dr.

INSTALACIJA PAKETA

```
# Instaliraj paket  
install.packages("lsr")  
# Učitaj paket  
library(lsr) # Također require()  
# Makni paket  
detach("package:lsr", unload=TRUE)
```

RADNI PROSTOR

```
# Gdje se nalazi radni folder  
getwd()  
[1] "C:/Users/Dropbox/Kolegij_statistika"  
# Postavi radni folder  
setwd("C:/Users/Luka/Statistika/...")  
# Obrati pažnju na "/" vs "\\"
```

UČITAVANJE U R I

- ▶ *workspace files* imaju .Rdata ekstenziju i sadržavaju objekte iz radnog prostora
- ▶ .csv, .txt, .spss sadržavaju podatke koje je potrebno učitati u radni prostor
- ▶ *script files* sadržavaju napisani kod
- ▶ postoje i brojni drugi objekti! (npr. RMarkdown)

```
# učitaj u radni prostor  
load( file = "~/datoteka.Rmd" )  
# ~ označava put do radnog direktorija
```

UČITAVANJE U R II

```
# Učitaj podatke iz excela  
podatci_statistika <- read.csv( file = "statistika.csv" )  
# U praksi je potrebno prilagoditi funkcijske argumente!
```

SPREMANJE

- ▶ obično želimo pohraniti: *radni prostor, output, skriptu*

```
# Pohrani radni prostor
```

```
save.image("~/Desktop/radni_prostor.RData")
```

```
# Pohrani podatke
```

```
save(neki_podatci, file = "neki_podatci.Rdata")
```

```
# Output je moguće pohraniti pomoću funkcije sink()
```


SPECIJALNE VRIJEDNOSTI

```
# Beskonačno
```

```
1 / 0
```

```
## [1] Inf
```

```
# Nije broj
```

```
0 / 0
```

```
## [1] NaN
```

- ▶ *nije dostupno* NA
- ▶ *nema vrijednosti* NULL

DAVANJE NAZIVA ELEMENTIMA VEKTORA I

```
# Stvori numerički vektor  
ocjena <- c( 1, 2, 3 )  
ocjena # Pregledaj objekt
```

```
## [1] 1 2 3
```

```
# Pripiši nazive  
names(ocjena) <- c("nedovoljan","dovoljan","dobar")  
ocjena # Pregledaj objekt
```

```
## nedovoljan    dovoljan    dobar  
##           1           2           3
```

DAVANJE NAZIVA ELEMENTIMA VEKTORA II

```
# Stvori numerički vektor sa nazivima v2.0  
ocjena <- c("nedovoljan" = 1, "dovoljan" = 2, "dobar" = 3)  
ocjena # Pregledaj objekt
```

```
## nedovoljan    dovoljan    dobar  
##           1           2           3
```

```
# Indeksiranje pomoću naziva  
ocjena["dobar"] # Pregledaj objekt
```

```
## dobar  
##      3
```

DAVANJE NAZIVA ELEMENTIMA VEKTORA III

```
# Vidi nazive svih elemenata u vektoru  
names(ocjena) # Pregledaj objekt
```

```
## [1] "nedovoljan" "dovoljan"    "dobar"
```

KLASE PODATAKA

```
x <- "tekst"    # x je tekst  
class(x)        # Pregledaj klasu
```

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

```
x <- TRUE       # x je logička klasa  
class(x)        # Pregledaj klasu
```

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

```
x <- 100        # x je broj  
class(x)        # Pregledaj klasu
```

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

FAKTORI I

```
# Stvori numerički vektor  
grupa <- c(1,1,1,2,2,3,3)  
grupa # Pogledaj objekt
```

```
## [1] 1 1 1 2 2 3 3
```

```
# Pretvori numerički vektor u faktor  
grupa <- as.factor(grupa)  
grupa # Pogledaj objekt
```

```
## [1] 1 1 1 2 2 3 3  
## Levels: 1 2 3
```

FAKTORI II

- ▶ računske operacije sa faktorima nisu iste kao sa numeričkim vektorima

```
grupa + 2
```

```
## Warning in Ops.factor(grupa, 2): '+' not meaningful for
```

```
## [1] NA NA NA NA NA NA NA
```

NAZIVI FAKTORA

```
# Stvori faktoraku varijablu; spol ispitanika  
spol <- as.factor(c(1,1,1,1,2,2,2))  
spol # Pogledaj objekt
```

```
## [1] 1 1 1 1 2 2 2  
## Levels: 1 2
```

```
# Pripiši nazive faktorima  
levels(spol) <- c("muški", "ženski")  
print(spol) # Prikaži objekt
```

```
## [1] muški muški muški muški ženski ženski ženski  
## Levels: muški ženski
```


PODATKOVNI OKVIRI (data frame)

```
# Za stvaranje df-a ćemo koristiti varijable spol i grupa
# Stvaranje dodatne dvije numeričke varijable
starost <- c(17, 19, 21, 37, 18, 19, 47)
bodovi <- c(12, 15, 16, 14, 25, 21, 29)
# Stvori df
df <- data.frame( grupa, spol, starost, bodovi )
head(df, 5) # Pogledaj prvih 5 redova df-a
```

##		grupa	spol	starost	bodovi
## 1	1		muški	17	12
## 2	1		muški	19	15
## 3	1		muški	21	16
## 4	2		muški	37	14
## 5	2		ženski	18	25

IZVLAČENJE PODATAKA IZ DFa

```
# Pogledaj seriju bodova  
df$bodovi
```

```
## [1] 12 15 16 14 25 21 29
```

```
# Pogledaj nazive varijabli (kolona)  
names(df)
```

```
## [1] "grupa" "spol" "starost" "bodovi"
```

LISTE

```
# Stvori listu  
ana <- list(godine = 26,  
            student = TRUE,  
            roditelji = c("Marko","Mia"))  
ana # Pogledaj objekt
```

```
## $godine  
## [1] 26  
##  
## $student  
## [1] TRUE  
##  
## $roditelji  
## [1] "Marko" "Mia"
```

LISTE I

```
# Pristupi elementima liste  
ana$student
```

```
## [1] TRUE
```

```
ana$roditelji
```

```
## [1] "Marko" "Mia"
```

LISTE II

```
# Dodaj element u listu  
ana$mjesto_stanovanja <- "Zagreb"  
ana # Pogledaj objekt
```

```
## $godine  
## [1] 26  
##  
## $student  
## [1] TRUE  
##  
## $roditelji  
## [1] "Marko" "Mia"  
##  
## $mjesto_stanovanja  
## [1] "Zagreb"
```