

Prostorne analize u *open source* GIS okruženju: R+SAGA+Google kartografski servisi

20-24 jun 2011. Građevinski fakultet, Univerzitet u Beogradu

Uvod u R

Milan Kilibarda

kili@grf.bg.ac.rs

Sadržaj

- Uvod u R okruženje (1-35),
- Podaci u R-u (36-55),
- R jezik (55-64),
- R grafici (65-77),
- Regresija (78-97),
- Literatura (97-102),

Uvod u R okruženje

R

R je sistem za statističke proračune i grafike.

R je besplatan i open source.

Pored ostalog R je programski jezik, sa velikim potencijalom za kreiranje grafika, nudi interfejs ka drugim programskim jezicima i sl.

R koriste: matematičari, statističari, geografi, biolozi, ekonomisti, psiholozi ...

R

Strma kriva učenja.

Nudi napredne rutine iz mnogih oblasti, gotove funkcije, više nego komercijalna rešenja za mnoge oblasti.

Manipulacija nad podacima se prati korak po korak.
(ponovljivost, ponovljivost uz modifikacije)

GUI rešenja = podešavanja koja vode krajnjem rešenju.
Udobnost + zavisnost od prodavca.

Kako preuzeti R

R se može preuzeti sa internet stranice:
<http://cran.r-project.org/>

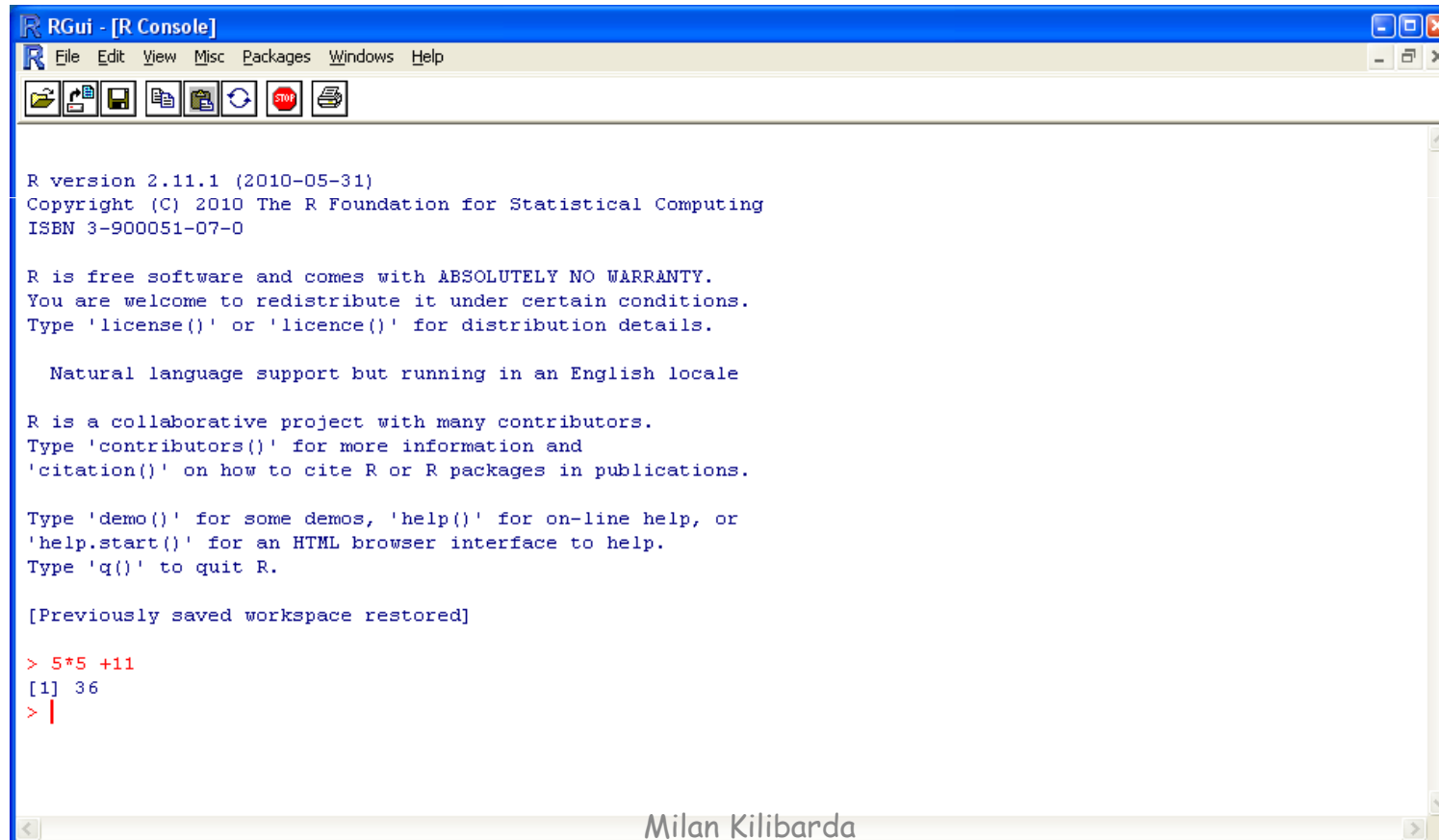
Instalacija zavisi od operativnog sistema.

Detaljna uputstva na srpskom jeziku:
http://savtajr.blogspot.com/2011/03/r-instalacija-jezika-i-paketa-sa-cran_11.html

Instalacija na Windows operativnom sistemu je krajnje jednostavna.

R konzola

Kad je instaliran R. Pokretanje konzole je isto kao i bilo kog drugog softvera.



```
RGui - [R Console]
File Edit View Misc Packages Windows Help

R version 2.11.1 (2010-05-31)
Copyright (C) 2010 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> 5*5 +11
[1] 36
> |
```

Milan Kilibarda

R osnovna instalacija

Osnovna instalacija R sadrži set paketa koje omogućavaju linearnu algebru, deskriptivnu statistiku, kreiranje grafika i sl.

Paketi su skupovi funkcija, dokumentacionih fajlova i podataka uvezanih zajedno. Pakete prave R korisnici i stručnjaci iz razvojnog tima.

Trenutno ima preko 3000 paketa.

NPR:

sp paket - set klasa i metoda koje omogućavaju manipulaciju nad prostornim podacima u R-u.

R paketi

Korisnici R-a instaliraju pakete koje čine njihovo interesno područje.

Korisnici koji se bave prostornim podacima instaliraju pakete koje sadrže funkcije za rad nad prostornim podacima. Kao i druge pakete koji im olakšavaju rad.

NPR:

sp, rgdal, gstat, maptools ...

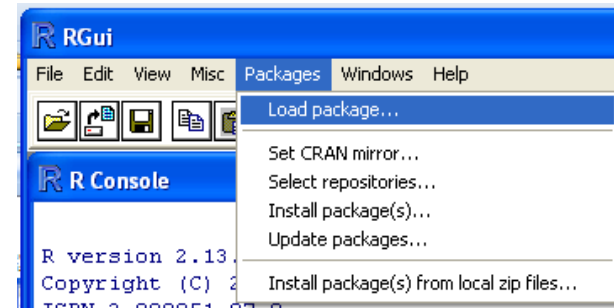
Instalacija R paketa

Detaljna lista paketa sa opisom dostupna je na stranici:

<http://cran.r-project.org/web/packages/>

Paket se može instalirati direktno iz konzole zadavanjem komande:

[install.packages](#)("plotGoogleMaps")



Instalacija više paketa

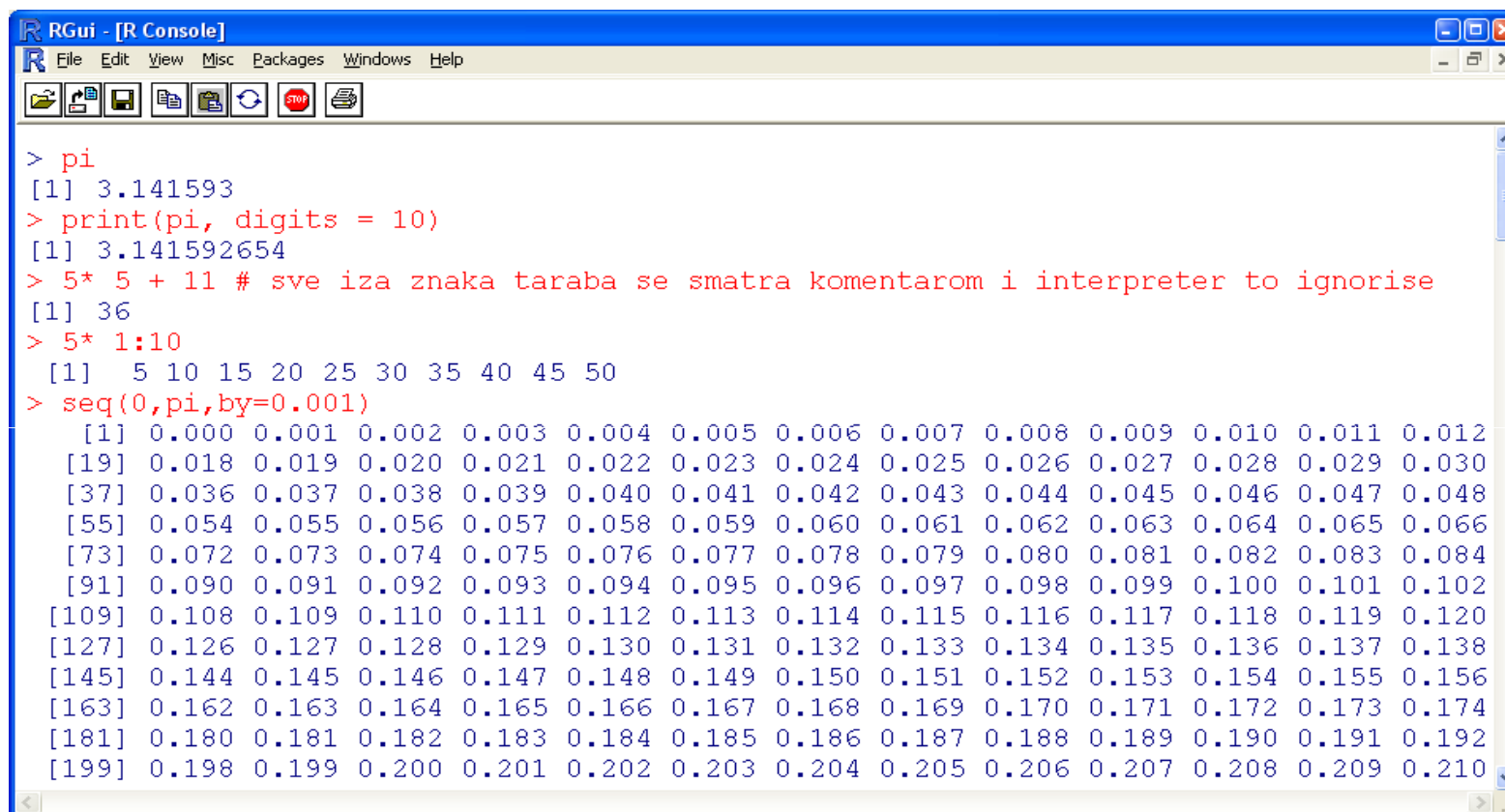
Instalacija više paketa odjednom:

```
install.packages(c("rgdal","maptools","gstat","spdep"  
,"spatstat","RSAGA","spgrass6"))
```

CRAN Task View - paketi su tematski grupisani.
Primer instalacije paketa iz Spatial grupe.

```
install.packages("ctv")  
library(ctv)  
install.views("Spatial")
```

R kao kalkulator



```
RGui - [R Console]
File Edit View Misc Packages Windows Help

> pi
[1] 3.141593
> print(pi, digits = 10)
[1] 3.141592654
> 5* 5+ 11 # sve iza znaka taraba se smatra komentarom i interpreter to ignorise
[1] 36
> 5* 1:10
[1] 5 10 15 20 25 30 35 40 45 50
> seq(0,pi,by=0.001)
[1] 0.000 0.001 0.002 0.003 0.004 0.005 0.006 0.007 0.008 0.009 0.010 0.011 0.012
[19] 0.018 0.019 0.020 0.021 0.022 0.023 0.024 0.025 0.026 0.027 0.028 0.029 0.030
[37] 0.036 0.037 0.038 0.039 0.040 0.041 0.042 0.043 0.044 0.045 0.046 0.047 0.048
[55] 0.054 0.055 0.056 0.057 0.058 0.059 0.060 0.061 0.062 0.063 0.064 0.065 0.066
[73] 0.072 0.073 0.074 0.075 0.076 0.077 0.078 0.079 0.080 0.081 0.082 0.083 0.084
[91] 0.090 0.091 0.092 0.093 0.094 0.095 0.096 0.097 0.098 0.099 0.100 0.101 0.102
[109] 0.108 0.109 0.110 0.111 0.112 0.113 0.114 0.115 0.116 0.117 0.118 0.119 0.120
[127] 0.126 0.127 0.128 0.129 0.130 0.131 0.132 0.133 0.134 0.135 0.136 0.137 0.138
[145] 0.144 0.145 0.146 0.147 0.148 0.149 0.150 0.151 0.152 0.153 0.154 0.155 0.156
[163] 0.162 0.163 0.164 0.165 0.166 0.167 0.168 0.169 0.170 0.171 0.172 0.173 0.174
[181] 0.180 0.181 0.182 0.183 0.184 0.185 0.186 0.187 0.188 0.189 0.190 0.191 0.192
[199] 0.198 0.199 0.200 0.201 0.202 0.203 0.204 0.205 0.206 0.207 0.208 0.209 0.210
```

Najčešće korišćeni operatori

$<-$	Dodeljivanje
$+$	Suma
$-$	Razlika
$*$	Množenje
$/$	Deljenje
$^$	Eksponent
$\% \%$	Moduo
$\% * \%$	Matrično množenje
$\% / \%$	Celobrojno deljenje

$ $	Ili
$\&$	i
$<$	Manje
$>$	Veće
$<=$	Manje ili =
$>=$	Veće ili =
$!$	Ne
$!=$	Različito
$==$	Isto

Operatori

> 2³

[1] 8

Koliko puta se broj 7 sadrži u broju 33 ?

> 33~~%~~/~~%~~7

[1] 4

Koliki je ostatak deljenja 33 sa 7 ?

> 33~~%~~%7

[1] 5

Promenljive

Sadrže vrednosti koje program smešta u memoriju računara.

Svaka promenljiva ima ime.

U R-u nije neophodna da se vrši deklaracija promenljive.

Imena promenljivih

R je CASE
SENSITIVE

Imena promenljivih

- Prom
- prom
- Milan11
- A_b
- A_matrica

LOŠE
promenljive

- ~~11Milan~~
- ~~A matrica~~
- ~~A*b~~
- ~~Prom#~~
- ~~print~~

Global environment

Global environment je radno okruženje u R-u gde mogu da se čuvaju rezultati računanja i mnogi drugi objekti.

`g <- 9.81 # konstanta je sačuvana u promenljivoj g`

`g2 <- g*2 # rezultat je sačuvana u promenljivoj g2`

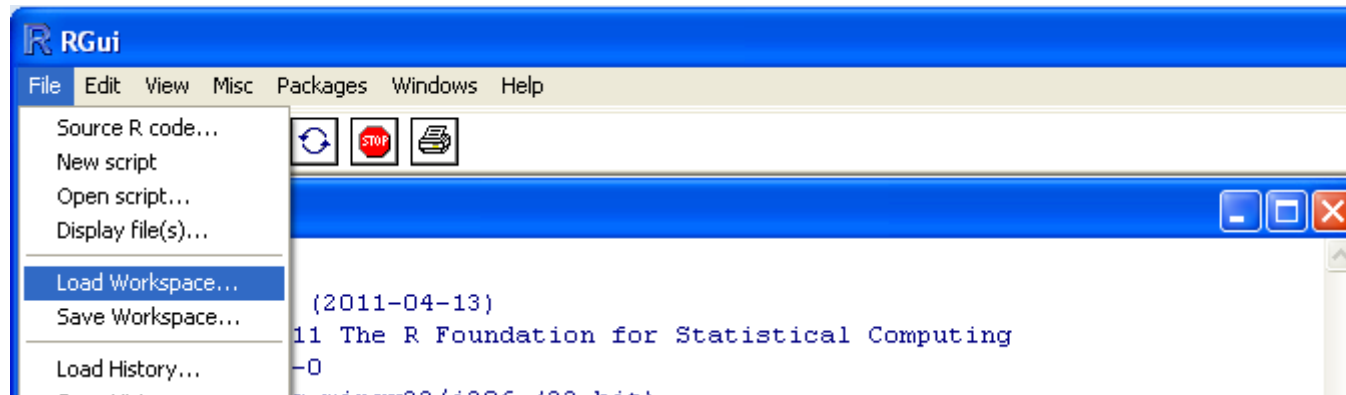
Okruženje se može sačuvati na više načina.

`save.image`("C:\\milan_r\\prvi_projekat.RData")

Global environment 2

Korišćenje prethodno sačuvanog okruženja:

load("C:\\milan_r\\prvi_projekat.RData")



Pregled i brisanje promeljivih u okruženju

```
ls() # lista svih promenljivih iz okruženja  
[1] "g" "g2"
```

```
rm("g") # brisanje prom. g iz okruž.
```

```
ls()
```

```
rm( list=ls() ) # brisanje svih prom. iz okruž.
```

```
(remove(almost)) # brisanje svih prom. iz okruž.
```

Ugradjene funkcije

R ima ogroman broj ugradjenih funkcija.
Neke su već korišćene do sad.

? load

load {base} R Documentation **Reload Saved Datasets**

Description

Reload datasets written with the function save.

Usage

load(file, envir = parent.frame()) **Arguments**

file a (readable binary-mode) [connection](#) or a character string giving the name of the file to load
(when [tilde expansion](#) is done).envir the environment where the data should be loaded. **Details**

load can load R objects saved in the current or any earlier format. It can read a compressed file
(see [save](#)) directly from a file or from a suitable connection (including a call to [url](#)).

A not-open connection will be opened in mode "rb" and closed after ...

...

Ugradjene funkcije

```
x<- c(1,2,3,4,8,11,18)
```

```
mean(x)
```

```
[1] 6.714286
```

```
? mean
```

```
help(mean)
```

Ugradjeni primeri

```
example(mean)
```

```
mean> x <- c(0:10, 50)
```

```
mean> xm <- mean(x)
```

```
mean> c(xm, mean(x, trim = 0.10))
```

```
[1] 8.75 5.50
```

```
mean> mean(USArrests, trim = 0.2)
```

```
  Murder  Assault UrbanPop  Rape
```

```
  7.42   167.60   66.20   20.16
```

Ako je definisano trim=0.1, onda se iz uzorka zanameri 10% najmanjih i najviših vrednosti. Tj računa se sredina od 80% uzorka.

Pomoć za funkcije kojima ne znamo ime

[help.search](#)("CRS")

Help files with alias or concept or title matching 'CRS' using regular expression matching:

```
raster::isLonLat      Is this longitude/latitude data?
rgdal::CRS-class      Class "CRS" of coordinate reference system
                      arguments
rgdal::spTransform-methods
                      Methods for Function spTransform for map
                      projection and datum transformation in package
                      "rgdal"
sp::CRS-class         Class "CRS" of coordinate reference system
                      arguments
sp::is.projected       Sets or retrieves projection attributes on
                      classes extending SpatialData
```

Type '?PKG::FOO' to inspect entries 'PKG::FOO', or 'TYPE?PKG::FOO' for entries like 'PKG::FOO-TYPE'.

Pomoć za funkcije kojima ne znamo ime

RSiteSearch("spTransform")



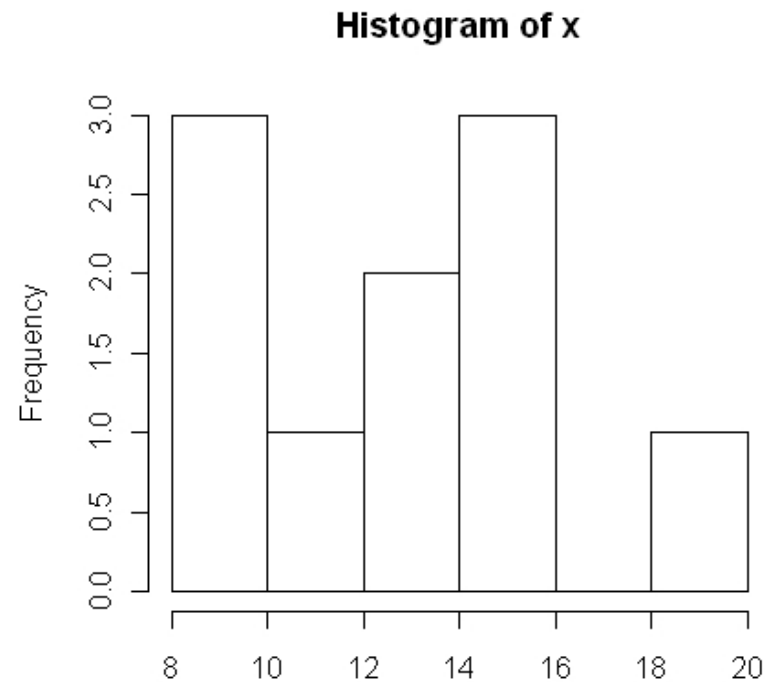
Najčešće korišćene f-je

c	Kobinacija više podataka
cbind, rbind	Po vrstama i kolonama
min	Minimum
max	Maksimum
length	# dužina (vek. i liste)
dim	# vrsta, kolona
floor	Zaokruživanje max.
which	Indeksi
table	Sumiranje po promenlj.

summary	Sumarna stat.
Sort, order, rank	Sortiranje ...
print	Prikaz u konzoli
cat	Prikaz kao karakteri
paste	c() spajanje karaktera
round	zaokruživanje
apply	Primeniti nešto na sve kolone ili vrste

Neke ugradjene grafičke f-je

```
x <- c(12, 15, 13, 20, 14, 16, 10, 10, 8, 15)  
hist(x)
```

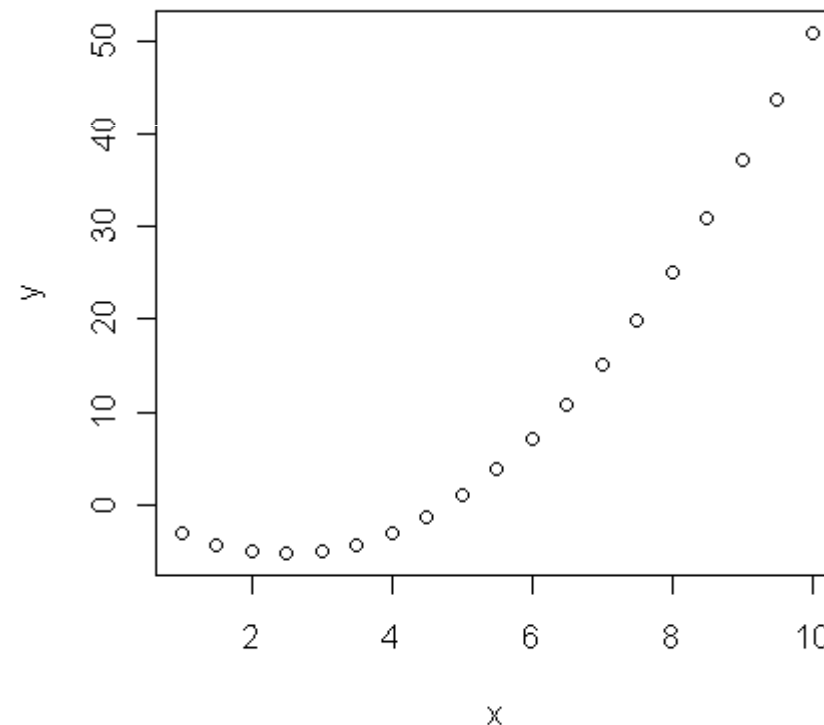


Neke ugradjene grafičke f-je

```
x <- seq(1, 10,by=0.5)
```

```
y<- x^2- 5*x +1
```

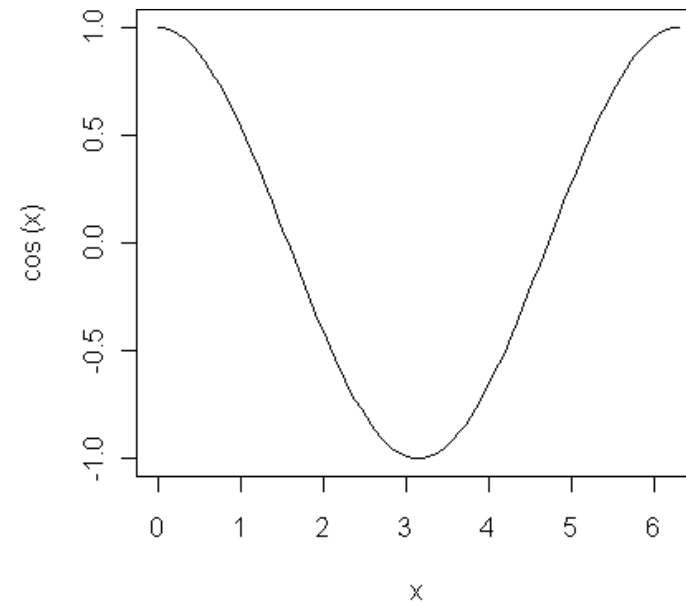
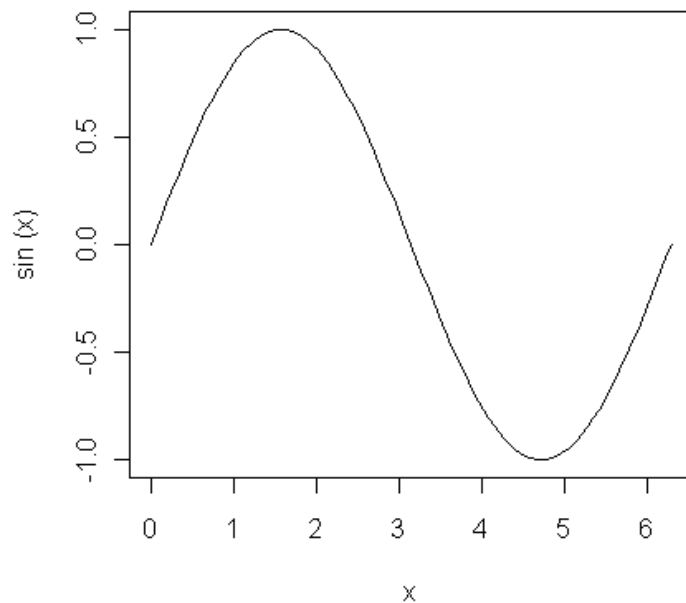
```
plot(x,y)
```



Neke ugradjene grafičke f-je

curve (expr = sin, from = 0, to=2*pi)

curve (expr = cos, from = 0, to=2*pi)



Podešavanje prikaza u konzoli

```
options( prompt="kili > ", continue="nastavk  
komande + ", digits=12, width=70,  
show.signif.stars=T)
```

```
kili > x
```

```
[1] 1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5 5.0 5.5 6.0  
6.5 7.0
```

```
[14] 7.5 8.0 8.5 9.0 9.5 10.0
```

Podešavanje jezika

Kako pruzeti podešavanja jezika iz OS (Windows)

Sys.getlocale()

```
[1] "LC_COLLATE=Serbian (Latin)_Serbia and  
Montenegro.1250;LC_CTYPE=Serbian (Latin)_Serbia and  
Montenegro.1250;LC_MONETARY=Serbian (Latin)_Serbia and  
Montenegro.1250;LC_NUMERIC=C;LC_TIME=Serbian (Latin)_Serbia and  
Montenegro.1250"
```

Sys.getlocale("LC_CTYPE")

```
[1] "Serbian (Latin)_Serbia and Montenegro.1250"
```

localeToCharset(locale = Sys.getlocale("LC_CTYPE"))

Rprofile.site

C:\Program Files\R\R-2.13.0\etc

Modifikacijom fajla Rprofile.site, mogu se izvršiti neka podešavanja koja će biti aktivna prilikom svakog pokretanja R-a.

Npr: podešavanje default paketa:

```
# Things you might want to change
```

```
# default packages
```

```
local({
```

```
old <- getOption("defaultPackages")
```

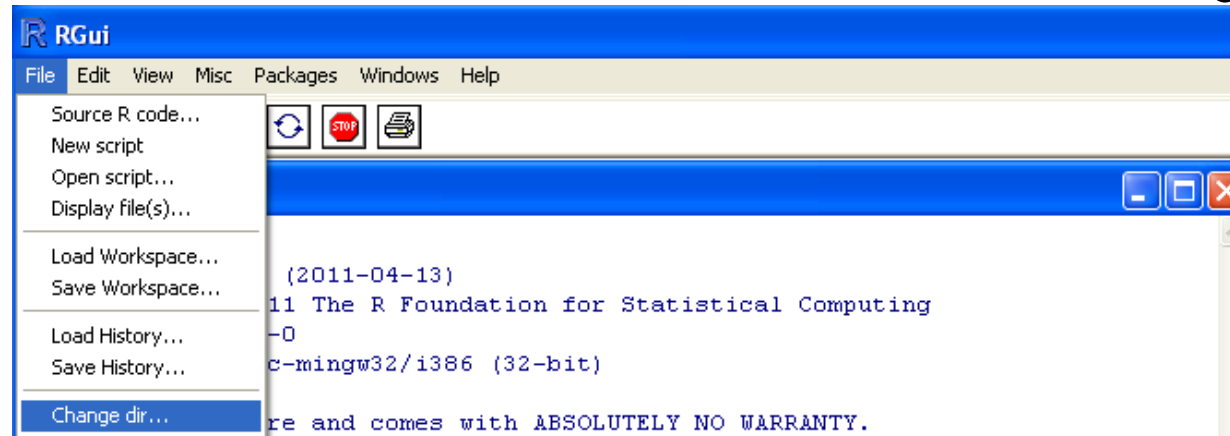
```
options(defaultPackages = c(old,"RSAGA", "gstat", "plotGoogleMaps"))
```

```
})
```

```
# proxy setting
```

```
Sys.setenv(http_proxy="http://mkkbbvzh:8080/")
```

Radni direktorijum



```
setwd("C:\\milan_r")
```

```
setwd('C:\\milan_r')
```

```
setwd("C:/milan_r")
```

```
setwd("C:\\milan_r")
```

```
getwd()
```


Editori

Može se koristiti bilo koji tekst editor,
kao što je NotePad i sl.

Preporučeni editori

Windows

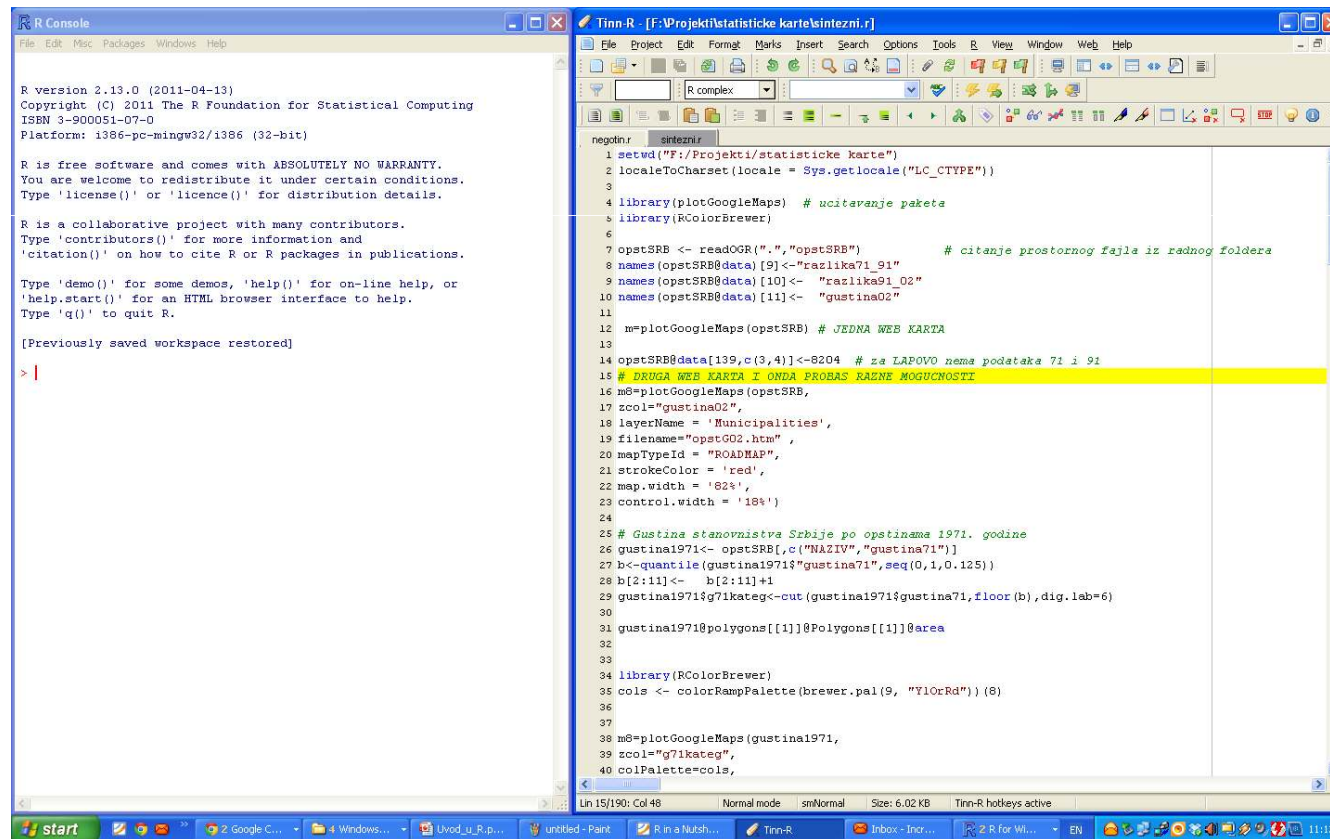
1. Tinn-R
2. Rstudio
3. Notepad+ with npptor
4. JGR

Linux

1. Rkward
2. Emacs+ESS

Tinn-R

<http://sourceforge.net/projects/tinn-r/>



The screenshot displays the Tinn-R environment. On the left, the R Console shows the standard R startup messages, including the version (2.13.0), copyright information, and platform details (i386-pc-mingw32/i386 (32-bit)). It also provides instructions on how to use the help system and how to cite R. The main window on the right contains an R script titled 'sintezni.r'. The script sets the working directory to 'F:/Projekti/statisticke_karte/sintezni.r', loads the 'plotGoogleMaps' and 'RColorBrewer' libraries, and reads a shapefile 'opstSRB'. It then creates a map of municipalities in Croatia, with a legend and a color scale. The script is written in R and includes comments in Croatian. The status bar at the bottom indicates the current line and column (Ln 15/190: Col 48) and the file size (6.02 KB).

```
R version 2.13.0 (2011-04-13)
Copyright (C) 2011 The R Foundation for Statistical Computing
ISBN 3-900051-07-0
Platform: i386-pc-mingw32/i386 (32-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

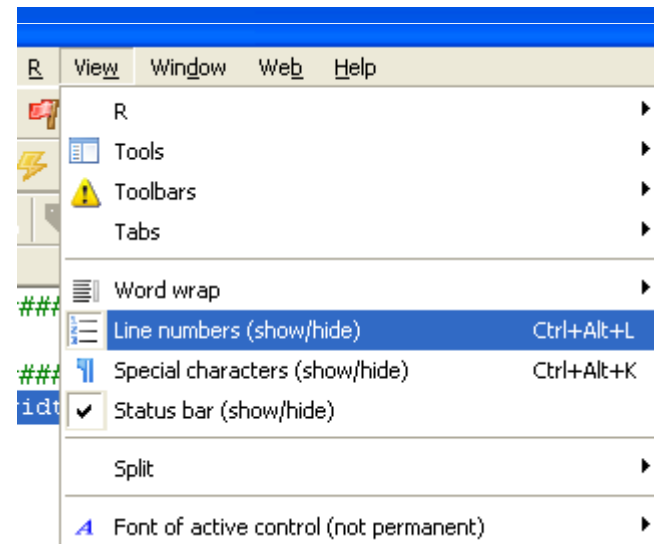
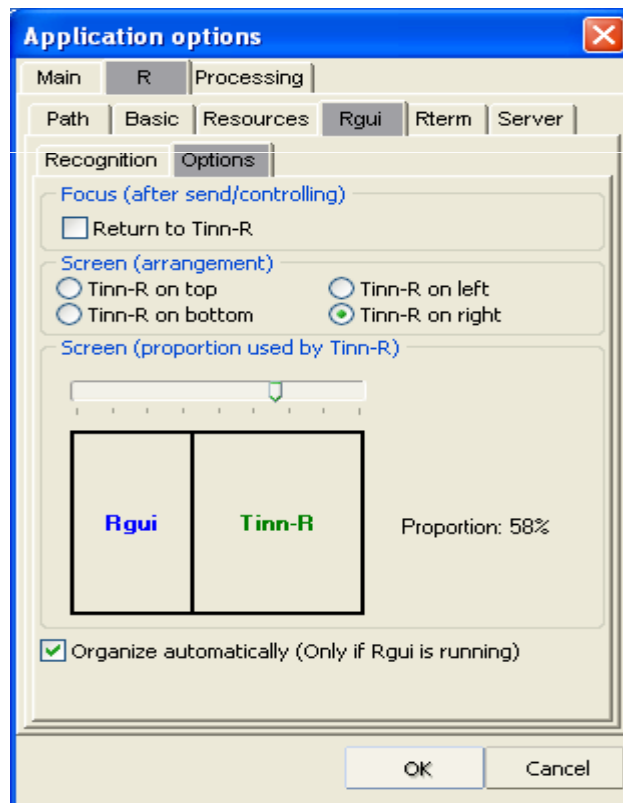
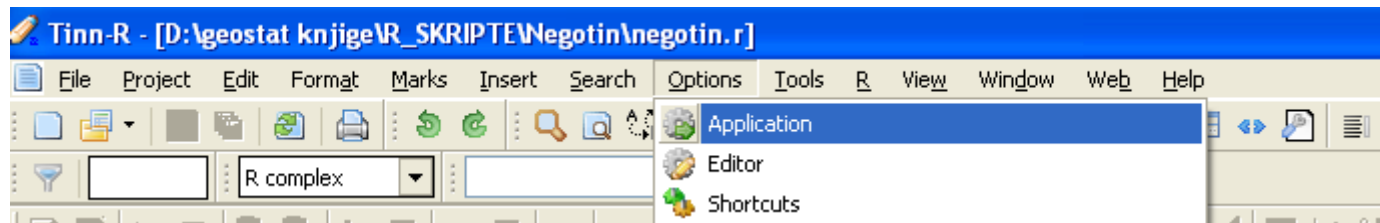
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]
> |
```

```
1 setwd("F:/Projekti/statisticke_karte")
2 localeToCharset(locale = Sys.getlocale("LC_CTYPE"))
3
4 library(plotGoogleMaps) # učitavanje paketa
5 library(RColorBrewer)
6
7 opstSRB <- readOGR(".", "opstSRB") # citanje prostornog fajla iz radnog foldera
8 names(opstSRB@data)[9] <- "razlika71_91"
9 names(opstSRB@data)[10] <- "razlika91_02"
10 names(opstSRB@data)[11] <- "gustina02"
11
12 m=plotGoogleMaps(opstSRB) # JEDNA WEB KARTA
13
14 opstSRB@data[139,c(3,4)] <- S204 # za LAPOVO nema podataka 71 i 91
15 # DRUGA WEB KARTA I ONDA PROBAS RAENE MOGUĆNOSTI
16 m8=plotGoogleMaps(opstSRB,
17   zcol="gustina02",
18   layerName = 'Municipalities',
19   fileName="opstG02.htm",
20   mapTypeId = "ROADMAP",
21   strokeColor = 'red',
22   map.width = '824',
23   control.width = '184')
24
25 # Gustina stanovništva Srbije po opstinama 1971. godine
26 gustina1971 <- opstSRB@c("NA2IV", "gustina71")
27 b <- quantile(gustina1971$gustina71, seq(0, 1, 0.125))
28 b[2:11] <- b[2:11]+1
29 gustina1971$g71kateg <- cut(gustina1971$gustina71, floor(b), dig.lab=6)
30
31 gustina1971@polygons[[1]]@Polygons[[1]]@area
32
33
34 library(RColorBrewer)
35 cols <- colorRampPalette(brewer.pal(9, "YlOrRd")) (8)
36
37
38 m8=plotGoogleMaps(gustina1971,
39   zcol="g71kateg",
40   colPalette=cols,
```

Tinn-R



Podaci u R-u

mode() funkcija

Tri osnovna tipa podataka (mod podataka)
u R-u su: **numerički, karakter i logički**
tip (numeric, character i logical)

Ovi podaci mogu da budu deo različitih
objekata (klasa) u R-u.

```
a<-c(1,2,3)  
mode(a)  
[1] "numeric"
```

```
b<-c("a","b","c")  
mode(b)  
[1] "character"
```

```
c<-c(TRUE,FALSE,TRUE)  
mode(c)  
[1] "logical"
```

NA: not available

Bilo koji tip podataka(numeric, character, logical) može sadržati NA: not available.

- NA nije 0
- NA nije ""
- NA nije FALSE
- NA nije NULL

```
max(c(NA, 4, 7))  
[1] NA  
max(c(NA, 4, 7),na.rm=T)  
[1] 7
```

```
NA | TRUE  
[1] TRUE  
NA & TRUE  
[1] NA
```

class() f-ja

Funkcija koja pokazuje klasu objekta -
class

```
f<-factor(c("a","b","c","a","a","b"))
```

```
f
```

```
[1] a b c a a b
```

```
Levels: a b c
```

```
mode(f)
```

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

```
class(f)
```

```
[1] "factor"
```

Vektori

Vektor u R-u je skup podataka istog tipa:

```
a<-c(1,2,3)
```

```
a*2
```

```
[1] 2 4 6
```

```
b<-c("a","b","c")
```

```
b
```

```
[1] "a" "b" "c"
```

```
C<-c(TRUE,FALSE,TRUE)
```

```
C
```

```
[1] TRUE FALSE TRUE
```


Vektori

U opštem slučaju bilo koji podatak zapisan u R-u direktno je vektor dužine 1.

```
a<-c(1,2,3)
```

```
length(a) # dužina vektora
```

```
[1] 3
```

```
summary(a)
```

```
Min. 1st Qu. Median Mean 3rd Qu. Max.
```

```
1.0 1.5 2.0 2.0 2.5 3.0
```

```
a[2]*8 # element vektora
```

```
[1] 16
```

```
as.character(a) # konverzija u tekstualni vektor
```

```
[1] "1" "2" "3"
```

Matrice

Matrica pravougaona tabela podataka istog tipa.

$$A = \begin{vmatrix} 11 & 8 & 88 \\ 2 & 4 & 33 \\ 3 & 2 & 11 \end{vmatrix}$$

```
A<-cbind(c(11,2,3),c(8,4,2), c(88,33,11))
```

A

```
      [,1] [,2] [,3]  
[1,]  11   8  88  
[2,]   2   4  33  
[3,]   3   2  11
```

```
A<-rbind(c(11,8,88),c(2,4,33),c(3,2,11) )
```

A

```
      [,1] [,2] [,3]  
[1,]  11   8  88  
[2,]   2   4  33  
[3,]   3   2  11
```

Matrice

```
x.mat <- matrix(c(3, -1, 2, 2, 0, 3, -3, 6), ncol = 2)
```

```
x.mat
```

```
      [,1] [,2]  
[1,]    3    0  
[2,]   -1    3  
[3,]    2   -3  
[4,]    2    6
```

```
x.mat[3,2]
```

```
[1] -3
```

```
dimnames(x.mat) <- list(c("R1","R2","R3","R4"),  
                        c("C1","C2"))
```

```
x.mat
```

```
      C1 C2  
R1    3  0  
R2   -1  3  
R3    2 -3  
R4    2  6
```

```
x.mat["R3","C2"]
```

```
[1] -3
```

Redovi

Redovi se najčešće koriste kao višedimenzionalne matrice ili vektori.

```
h<-1:24
```

```
Z <- array(h, dim=c(3,4,2))
```

```
Z
```

```
.,1
```

```
      [,1][,2][,3][,4]  
[1,]  1  4  7 10  
[2,]  2  5  8 11  
[3,]  3  6  9 12
```

```
.,2
```

```
      [,1][,2][,3][,4]  
[1,] 13 16 19 22  
[2,] 14 17 20 23  
[3,] 15 18 21 24
```

Data Frame (tabela podataka)

Najčešće korišćeni tip podataka u R-u.

Data frame je tabela. Svaku kolonu čine podaci istog tipa.

Različite kolone mogu biti različitog tipa.

Najviše podseća na Excel tabelu.

Data Frame

```
student = c("Milan", "Jelena", "Marko")
```

```
god_stud = c(3, 3, 1)
```

```
ranije_progr = c(TRUE, FALSE, TRUE)
```

```
df = data.frame(student , god_stud, ranije_progr)
```

df

	student	god_stud	ranije_progr
1	Milan	3	TRUE
2	Jelena	3	FALSE
3	Marko	1	TRUE

```
fix(df)
```

Data Frame - selekcija

str(df)

```
'data.frame': 3 obs. of 3 variables:  
 $ student   : Factor w/ 3 levels  
   "Jelena","Marko",...: 3 1 2  
 $ god_stud   : num 3 3 1  
 $ ranije_progr: logi TRUE FALSE TRUE
```

df\$student

```
[1] Milan Jelena Marko  
Levels: Jelena Marko Milan
```

df[,1]

```
[1] Milan Jelena Marko  
Levels: Jelena Marko Milan
```

df[, 'god_stud']

```
[1] 3 3 1
```

df[df\$student == 'Milan',]

```
  student god_stud ranije_progr  
1  Milan      3      TRUE
```

df[df\$god_stud > 1,]

```
  student god_stud ranije_progr  
1  Milan      3      TRUE  
2  Jelena      3      FALSE
```

df[df\$god_stud > 1 & df\$ranije_progr == T,]

```
  student god_stud ranije_progr  
1  Milan      3      TRUE
```

Liste

Uredjen skup podataka koji ne moraju biti istog tipa. Npr: lista = [vektor,karaktri, df, lista]

```
kurs.l <- list(ime="R kurs", br_ucesnika=3, podaci=df)
```

```
$ime
```

```
[1] "R kurs"
```

```
$br_ucesnika
```

```
[1] 3
```

```
$podaci
```

```
  student god_stud ranije_progr  
1  Milan      3      TRUE  
2  Jelena      3     FALSE  
3   Marko      1      TRUE
```

```
kurs.l$ime
```

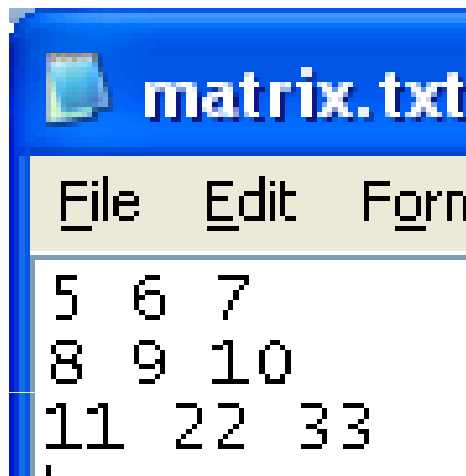
```
kurs.l[[3]]
```

```
kurs.l[[3]]$student
```


Ispitivanje podataka

class	Klasa objekta. (vector, matrix, function, logical, list, ...)
str	Struktura podataka
mode	Tip podataka. (Numeric, character, logical, ...)
storage.mode typeof	Tip koji koristi R da bi skladištio podatak u memoriji (double, integer, character, logical, ...)
is.function	Da li je funkcija(TRUE if function)
is.na	Da li je NA (TRUE if missing)
names	Imena pridružena objektu
dimnames	Imena za indekse kod vektora, matrica i redova
slotNames	Imena slotova - delova objekta (npr SP podaci)
attributes	Imena i klase atributa kod objekta...

Čitanje i pisanje numeričkih podataka u fajl



```
matrix.txt
File Edit Format
5 6 7
8 9 10
11 22 33
```

```
scan("matrix.txt")
```

```
Read 9 items
```

```
[1] 5 6 7 8 9 10 11 22 33
```

```
matrix(scan("matrix.txt"),ncol=3,byrow=TRUE)
```

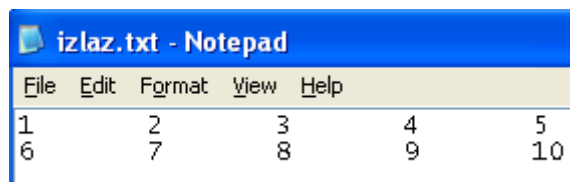
```
Read 9 items
```

```
      [,1][,2][,3]
[1,]  5   6   7
[2,]  8   9  10
[3,] 11  22  33
```

```
x <- matrix(1:10,ncol=5)
```

```
write(x, "izlaz1.txt", sep = "\t")
```

```
write(x, "izlaz1.data", sep = "\t")
```



```
izlaz.txt - Notepad
File Edit Format View Help
1      2      3      4      5
6      7      8      9     10
```

Čitanje i pisanje tabelarnih podataka

Osnovna f-ja za čitanje i pisanje tabele je `read.table`, `write.table`.

Učitana tabela je `data.frame` objekat.

Neke dodatne funkcije:

`read.csv`, `write.csv`

`read.csv2`, `write.csv2`

`read.delim`, `write.delim`

`read.delim2`, `write.delim2`

`read.csv` and `read.csv2` are identical to `read.table` except for the defaults. They are intended for reading 'comma separated value' files ('.csv') or (read.csv2) the variant used in countries that use a comma as decimal point and a semicolon as field separator.

Primer za tabele

ID	X	Y	Soil_type	Humus	pH_H2O	Total_N	K2O	P2O5	T_S	S	T	V	Sand	Silt	Clay	Coarse_fraction	Fine_fraction
1	7618924.72	4895559.46	6	2.73	5.1	0.12	25.3	9.1	11.9	13	24.9	52.2	39.9	33.3	26.8	61.3	38.7
2	7618870.98	4895816.88	6	2.94	6.52	0.13	19.2	2.2	4.4	27.3	31.7	86	32.4	29.3	38.3	51.8	48.2
3	7618825.72	4895989.43	7	2.35	5.64	0.11	16.9	7.9	8.1	16.2	24.3	66.7	51.6	21.3	27.1	65	35
4	7618818.65	4896221.39	7	2.47	5.96	0.11	23.3	6.5	6.3	18	24.3	68.9	42	25.7	32.3	59.7	40.3
5	7618916.24	4896662.67	6	2.17	5.9	0.09	15	1.7	5.6	15.3	20.9	73.2	47.3	25.4	27.3	63.9	36.1
6	7619080.3	4896395.35	2	3.64	7.7	0.2	32.2	16.3	1.3	57.9	59.2	97.7	30.4	31.8	37.8	50.4	49.6
7	7619277.89	4896024.79	2	3.36	7.14	0.2	32.2	22.5	2.9	33	35.9	92	33.3	29.7	37	50.8	49.2
8	7619746.33	4895922.95	4	3.03	6.84	0.14	22.7	2.4	3.5	24.1	27.6	87.3	37.1	30.1	32.8	56	44

```
profili <- read.delim("profiles.txt")
str(profili)
```

```
'data.frame': 151 obs. of 18 variables:
 $ ID      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ X       : num  7618925 7618871 7618826 7618819 7618916 ...
 $ Y       : num  4895559 4895817 4895989 4896221 4896663 ...
 $ Soil_type : int  6 6 7 7 6 2 2 4 4 2 ...
 $ Humus    : num  2.73 2.94 2.35 2.47 2.17 3.64 3.36 3.03 2.22 3.04
 ...
 $ pH_H2O   : num  5.1 6.52 5.64 5.96 5.9 7.7 7.14 6.84 6.36 6.44
 ...
 $ Total_N  : num  0.12 0.13 0.11 0.11 0.09 0.2 0.2 0.14 0.1 0.14 ...
 $ K2O      : num  25.3 19.2 16.9 23.3 15 32.2 32.2 22.7 21.2 23.6
 ...
 $ P2O5     : num  9.1 2.2 7.9 6.5 1.7 16.3 22.5 2.4 2.4 4.9 ...
 $ T_S      : num  11.9 4.4 8.1 6.3 5.6 1.3 2.9 3.5 4.6 5.5 ...
 $ S        : num  13 27.3 16.2 18 15.3 57.9 33 24.1 10.4 20 ...
 $ T        : num  24.9 31.7 24.3 24.3 20.9 59.2 35.9 27.6 15 25.5 ...
 $ V        : num  52.2 86 66.7 68.9 73.2 97.7 92 87.3 69.4 78.5 ...
 $ Sand     : num  39.9 32.4 51.6 42 47.3 30.4 33.3 37.1 60.1 51.4
 ...
 $ Silt     : num  33.3 29.3 21.3 25.7 25.4 31.8 29.7 30.1 22.1 23.2
 ...
 $ Clay     : num  26.8 38.3 27.1 32.3 27.3 37.8 37 32.8 17.8 25.4 ...
 $ Coarse_fraction: num  61.3 51.8 65 59.7 63.9 50.4 50.8 56 74.5 65.6 ...
 $ Fine_fraction : num  38.7 48.2 35 40.3 36.1 49.6 49.2 44 25.5 34.4 ...
```

Primer 2 za tabele

`write.csv(profil, "profil.csv")` # ili

`write.csv2(profil, "profil.csv")` # u zavisnosti šta je definisano kao separator za csv fajl

profil.csv - Microsoft Excel

Menus Home Insert Page Layout Formulas Data Review View

All File Edit View Insert Format Tools Data Window Help

Calibri 11 B I U General % 0.00 0.00

Toolbars

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T		
1	ID	X	Y	Soil_type	Humus	pH_H2O	Total_N	K2O	P2O5	T_S	S	T	V	Sand	Silt	Clay	Coarse_fr	Fine_fraction				
2	1	1	7618924,7	4895559,4	6	2,73	5,1	0,12	25,3	9,1	11,9		13	24,9	52,2	39,9	33,3	26,8	61,3	38,7		
3	2	2	7618870,9	4895816,8	6	2,94	6,52	0,13	19,2	2,2	4,4	27,3	31,7		86	32,4	29,3	38,3	51,8	48,2		
4	3	3	7618825,7	4895989,4	7	2,35	5,64	0,11	16,9	7,9	8,1	16,2	24,3	66,7	51,6	21,3	27,1		65	35		
5	4	4	7618818,6	4896221,3	7	2,47	5,96	0,11	23,3	6,5	6,3		18	24,3	68,9		42	25,7	32,3	59,7	40,3	
6	5	5	7618916,2	4896662,6	6	2,17	5,9	0,09		15	1,7	5,6	15,3	20,9	73,2	47,3	25,4	27,3	63,9	36,1		
7	6	6	7619080,3	4896395,3	2	3,64	7,7	0,2	32,2	16,3	1,3	57,9	59,2	97,7	30,4	31,8	37,8	50,4	49,6			
8	7	7	7619277,8	4896024,7	2	3,36	7,14	0,2	32,2	22,5	2,9		33	35,9		92	33,3	29,7		37	50,8	49,2
9	8	8	7619746,3	4895922,9	4	3,03	6,84	0,14	22,7	2,4	3,5	24,1	27,6		87,3	37,1	30,1	32,8		56	44	
10	9	9	7619971,4	4895927,4	4	2,22	6,36	0,1	21,2	2,4	4,6	10,4		15	69,4	60,1	22,1	17,8	74,5	25,5		
11	10	10	7620184,2	4895860,7	2	3,04	6,44	0,14	23,6	4,9	5,5		20	25,5	78,5	51,4	23,2	25,4	65,6	34,4		
12	11	11	7620298,4	4895884,4	3	3,58	7,08	0,18	24,5	4,7	3,6	20,9	24,5		85,1	58,9	20,1		21	72	28	
13	12	12	7619598,1	4896008,9	6	1,83	6,68	0,09	25,9	2,1	3,7	24,1	27,8		86,6	31,1	31,7	37,2	50,6	49,4		
14	13	13	7618602,2	4895730,6	8	1,93	5,7	0,1	10,2	5,8	8,4	9,5	17,8		53,1	46,1	33,6	20,3	68,4	31,6		

Ready Milan Kilibarda 100%

R i baze podataka

Slično R može direktno da manipuliše i sa bazama podataka, sa Excel fajlovima itd.

Pored toga postoje paketi koji omogućavaju direktnu manipulaciju sa prostornim bazama podataka kao što je **RODBC**, **RpostGIS** paketi.

R jezik

Kontrola toka - If

```
if (logički uslov) {  
    izrazi (statements)  
}
```

```
else {  
    alternativni izrazi  
}
```

{ } su opcione u slučaju samo jednog izraza

```
ifelse (logički uslov, da izraz, ne izraz)
```


If primer

```
x<-3
```

```
if(x>2) {
```

```
  y<-2*x } else {
```

```
  y<-3*x}
```

```
y<-ifelse(x>2, 2*x, 3*x)
```

Petlje

```
for(i in 1:10) {  
  print(i*i+2*i)  
}
```

```
i<-1  
while(i<=10) {  
  print(i*i)  
  i<-i+sqrt(i)  
}
```

Pogledati: repeat, break, next

Funkcije

Funkcijama kreiramo neku proceduru nad podacima, vrši se obrada ulaznih podataka i dobija rezultat.

“Input”: ulazni argumenti, jedan ili više

“Output”: rezultat jedan objekat

```
gs <- function(a,b) {  
  result <- sqrt(a*b)  
  return(result)  
}
```

```
gs(7,88)
```

```
[1] 24.8193472919817
```

lapply, sapply, apply

Ugrađene funkcije u R-u koji nam olakšavaju manipulaciju nad vektorima, listama i matricama.

Izbegava se kontrola elementa **koristeći petlje.**

apply

```
A<-cbind(c(11,2,3) , c(8,4,2) , c(88,33,11) )
```

```
[,1][,2][,3]
```

```
[1,] 11 8 88
```

```
[2,] 2 4 33
```

```
[3,] 3 2 11
```

```
apply(A, 1, sum) # suma po vrstama
```

```
[1] 107 39 16
```

```
apply(A, 2, sum) # suma po kolonama
```

```
[1] 16 14 132
```

```
apply(A, 2, sort) # sortiranje po kolonama
```

```
[,1][,2][,3]
```

```
[1,] 2 2 11
```

```
[2,] 3 4 33
```

```
[3,] 11 8 88
```

apply

apply(A,1, function(x) x/max(x))

	[,1]	[,2]	[,3]	
[1,]	0.12500000	0.06060606	0.2727273	11/88
[2,]	0.09090909	0.12121212	0.1818182	8 /88
[3,]	1.00000000	1.00000000	1.0000000	88/88

$$A = \begin{vmatrix} 11 & 8 & 88 \\ 2 & 4 & 33 \\ 3 & 2 & 11 \end{vmatrix}$$

apply(A,2, function(x) x/max(x))

	[,1]	[,2]	[,3]	
[1,]	1.0000000	1.00	1.000	11/11
[2,]	0.1818182	0.50	0.375	2/11
[3,]	0.2727273	0.25	0.125	3/11

lapply, sapply


```
numList<-list(x=1:10,y=rnorm(100, 5, 5), z=rpois(100, 45))
```

```
lapply(numList, mean)
```

```
sapply(numList, mean)
```



```
x      y      z  
5.500000 5.595398 44.720000
```



```
$x  
[1] 5.5  
  
$y  
[1] 5.595398  
  
$z  
[1] 44.72
```

```
txt<-c("GRF","stat","Rcourse")
```

```
lapply(txt, nchar)
```

```
sapply(txt, nchar)
```



```
IBS  stat Rcourse  
3    4    7
```



```
[[1]]  
[1] 3  
  
[[2]]  
[1] 4  
  
[[3]]  
[1] 7
```

supply

```
fajlovi_wd<- list.files()  # dir() je isto kao list.files()
```

```
fajlovi_wd
```

```
[1] "izlaz.txt" "matrix.txt" "myMap.htm" "OGS_2010.pdf" "profili.csv"
```

```
a<- regexpr(".txt",fajlovi_wd[[1]],fixed=T)
```

```
> a
```

```
[1] 6
```

```
attr(,"match.length")
```

```
[1] 4
```

```
attr(a,"match.length")
```

```
[1] 4
```

```
vec<-sapply(fajlovi_wd,function(x) attr(regexpr(".txt",x,fixed=T),  
  "match.length") ==4)
```

```
vec
```

izlaz.txt	matrix.txt	myMap.htm	OGS_2010.pdf	profili.csv
TRUE	TRUE	FALSE	FALSE	FALSE

```
fajlovi_wd[vec]
```

```
[1] "izlaz.txt" "matrix.txt"
```


R grafici

Podaci

Podaci korišćeni u ovom delu su podaci ugrađeni u R.

`data()`

`data(cars)`

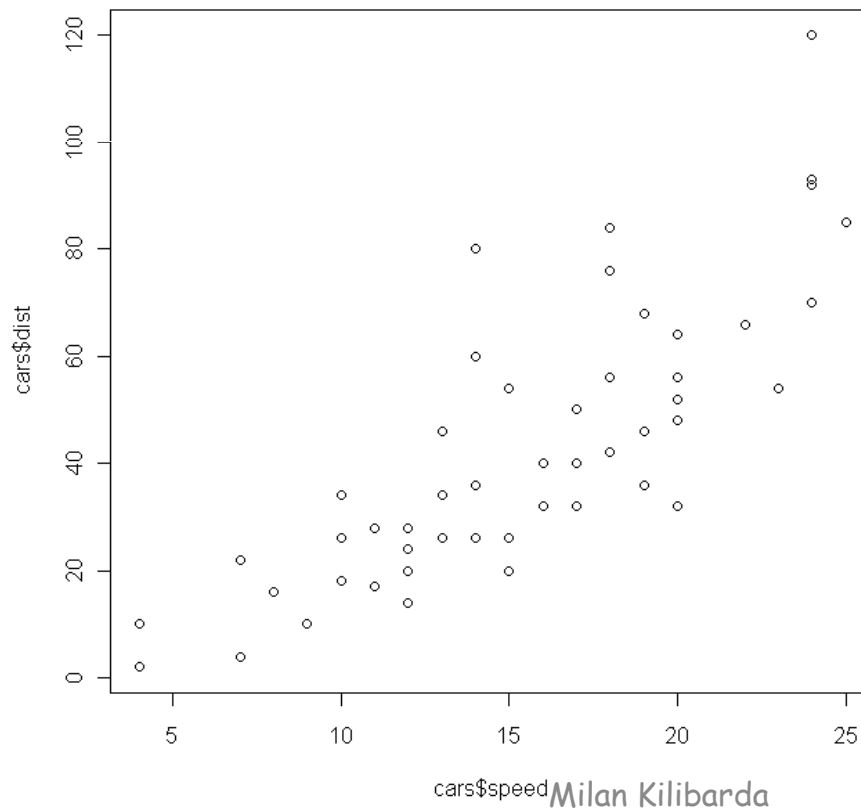
`help(cars)`

`str(cars)`

Podaci cars sadrže brzinu automobila i zaustavni put. Podaci su iz 1920. god.

plot

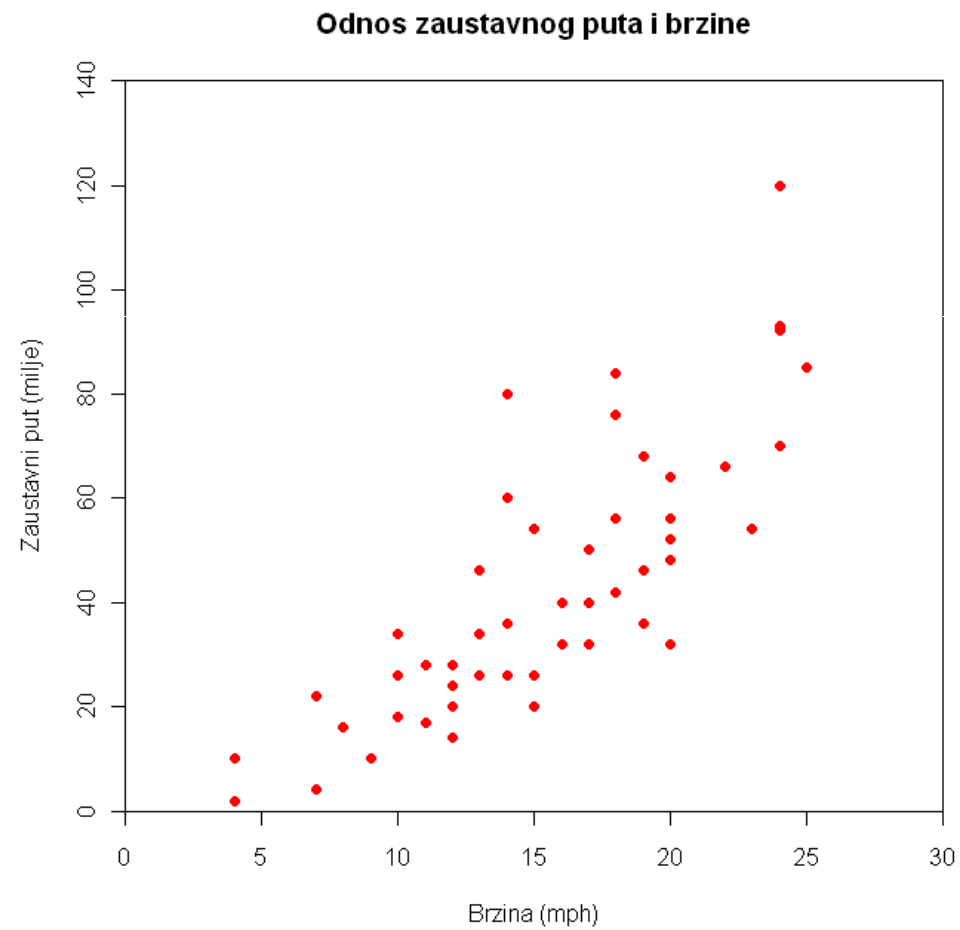
```
plot(cars$dist~cars$speed)
```



plot

```
plot(cars$dist~cars$speed, # y~x  
main='Odnos zaustavnog puta i brzine', # Naslov  
xlab='Brzina (mph)', # X osa naslov  
ylab='Zaustavni put (milje)', # Y osa naslov  
xlim=c(0,30), # X osa u granicama 0-30  
ylim=c(0,140), # Y osa u granicama 0-30  
xaxs='i', # internal stil X ose  
yaxs='i', # internal stil Y ose  
col='red', # Boja plotting simbola  
pch=19) # Velicina simbola u tackama
```

plot



barplot

```
barplot(c(h1, h2, h3))
```

```
data(airquality)
```

```
str(airquality)
```

```
heights <- tapply(airquality$Temp, airquality$Month, mean)
```

```
barplot(heights)
```

```
barplot(heights,
```

```
  main="Mean Temp. by Month",
```

```
  names.arg=c("May", "Jun", "Jul", "Aug", "Sep"),
```

```
  ylab="Temp (deg. F) ",
```

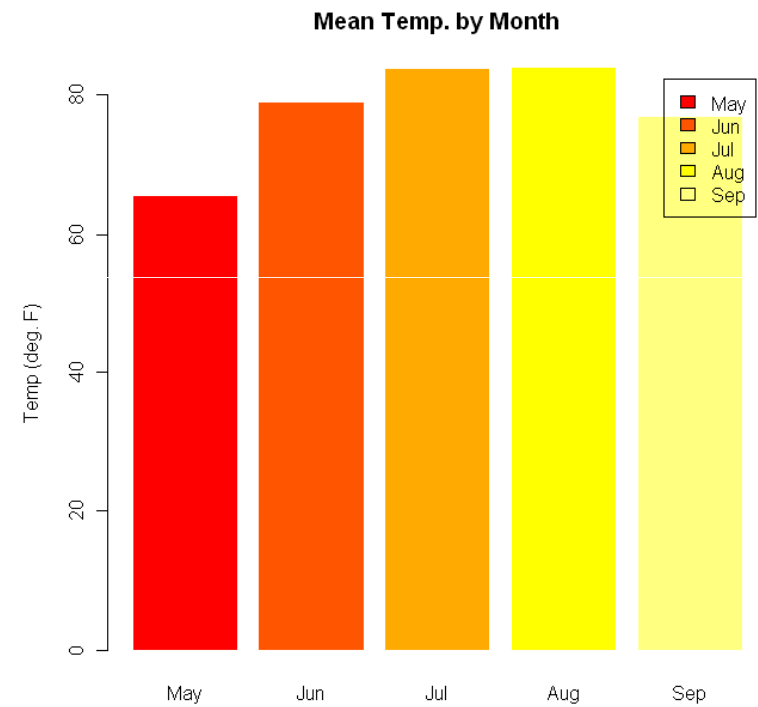
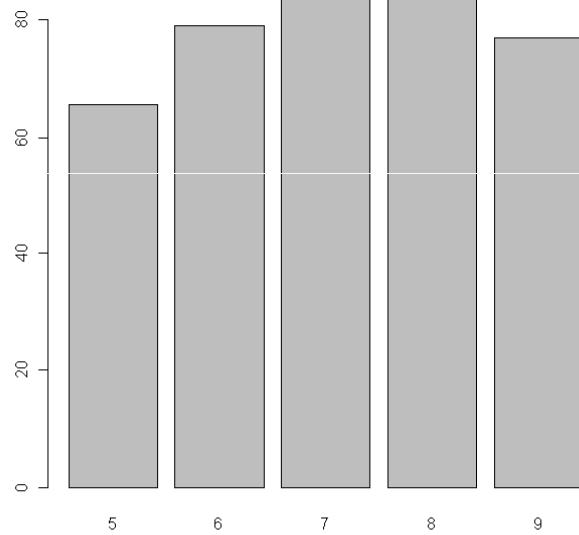
```
  legend=c("May", "Jun", "Jul", "Aug", "Sep"),
```

```
  col=heat.colors(5),
```

```
  border="white")
```

Milan Kilibarda

barplot



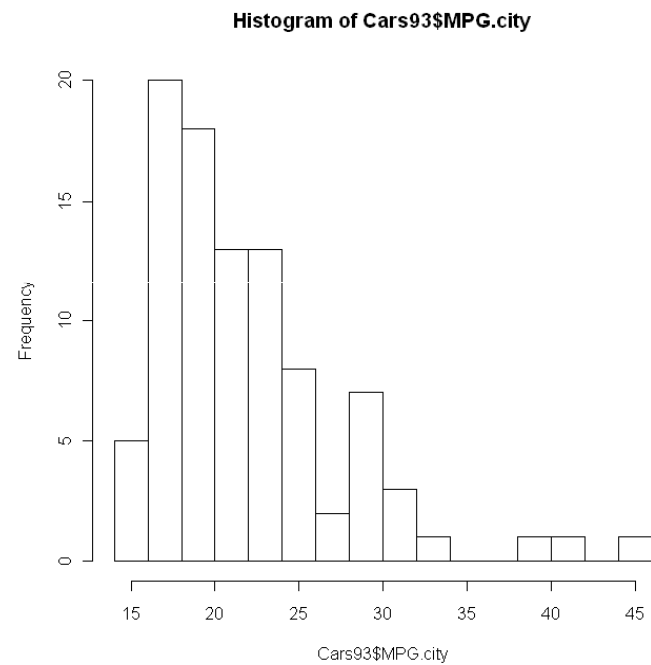
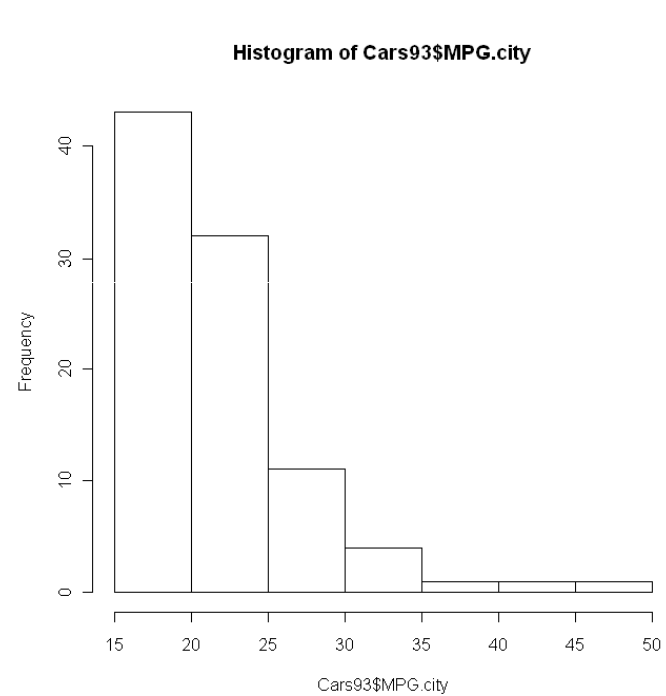
hist

```
data(Cars93, package="MASS")  
str(Cars93)
```

```
hist(Cars93$MPG.city)
```

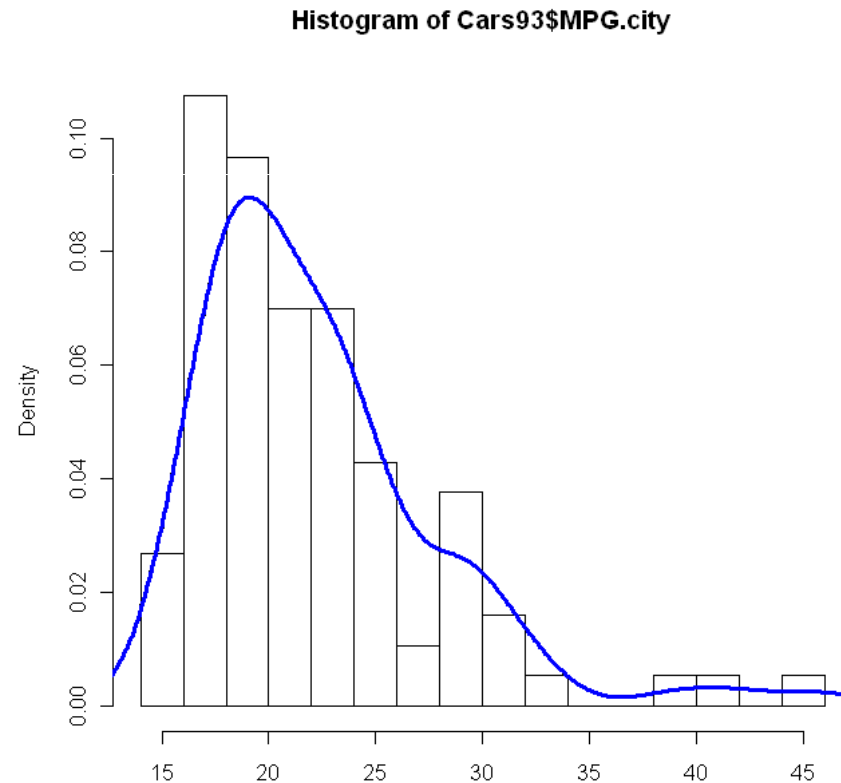
```
hist(Cars93$MPG.city,20)  
locator()
```


hist



Histogram sa funkcijom gustine raspodele

```
dest <- density(Cars93$MPG.city)  
hist(Cars93$MPG.city,20,probability = TRUE)  
lines(dest, lty=2)
```



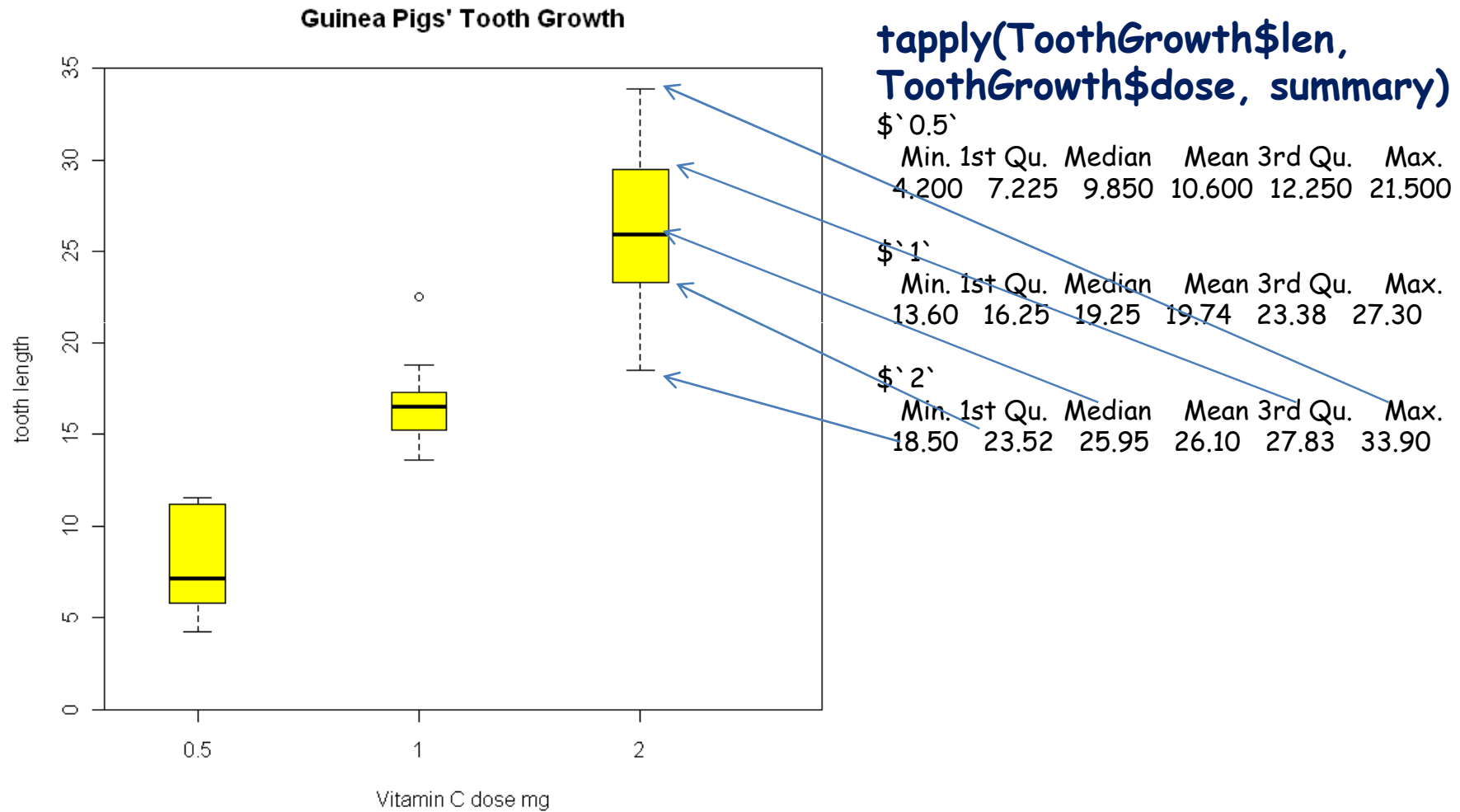
boxplot

```
data(ToothGrowth)  
str(ToothGrowth)
```

```
'data.frame': 60 obs. of 3 variables:  
 $ len: num 4.2 11.5 7.3 5.8 6.4 10 11.2 11.2 5.2 7 ...  
 $ supp: Factor w/ 2 levels "OJ","VC": 2 2 2 2 2 2 2 2 2 2 ...  
 $ dose: num 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
```

```
boxplot(len ~ dose, data = ToothGrowth, boxwex = 0.25, at =  
 1:3 - 0.2, subset = supp == "VC", col = "yellow", main =  
 "Guinea Pigs' Tooth Growth", xlab = "Vitamin C dose mg",  
 ylab = "tooth length", xlim = c(0.5, 3.5), ylim = c(0, 35),  
 yaxs = "i")
```

boxplot



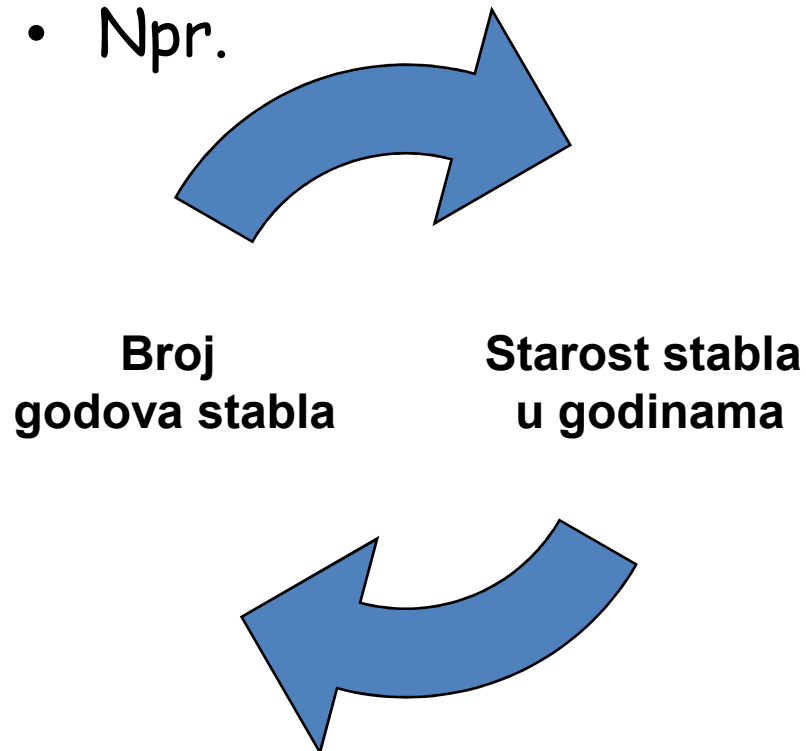
Neke funkcije za grafike

plot	Generički za grafike ...
points	Dodaje tačke
lines, abline	Dodaje linije
text, mtext	Dodaje tekst
legend	Legendu
axis	Koord. ose
box	Okvir i ose
par	Parametri: debljina linije, boja, tip linije , tacke
colors, palette	Boje

Regresija

Prosta linearna regresija

- Funkcionalana zavisnost ?
- Npr.



X - broj godova stabla

Y - starost stabla

Funkcionalana zavisnost

$$Y = f(X)$$

$$Y = X$$

Ako znamo broj godova
tada tacno znamo i
starost stabla.

Prosta linearna regresija

- Stohastička zavisnost
 - Ne može se utvrditi funkcionalna zavisnost;
 - Eksperimentalni skup podataka izmerenih vrednosti parova x i Y ;
 - Na ordinati se nanose vrednosti nezavisno promenljive x , a na apcisi zavisno promenljive Y (koja je slučajna promenljiva);
 - Rezultat je dijagram rasipanja, grafički prikaz zavisnosti i međuzavisnosti između x i Y .

Prosta linearna regresija

- Stohastička zavisnost?

Npr

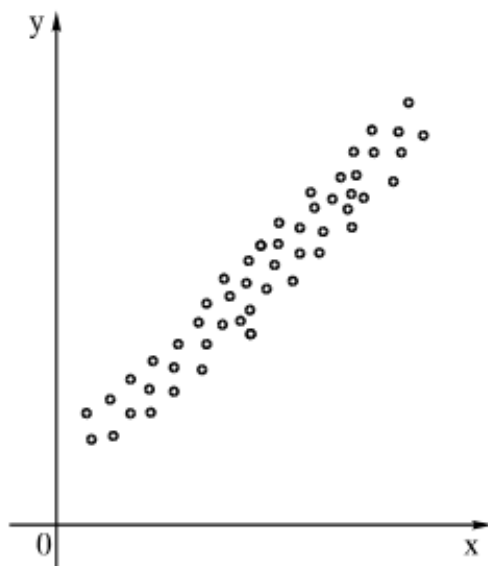
x - broj sati koje je osoba W provela trenirajući;

Y - telesna težina osobe W .

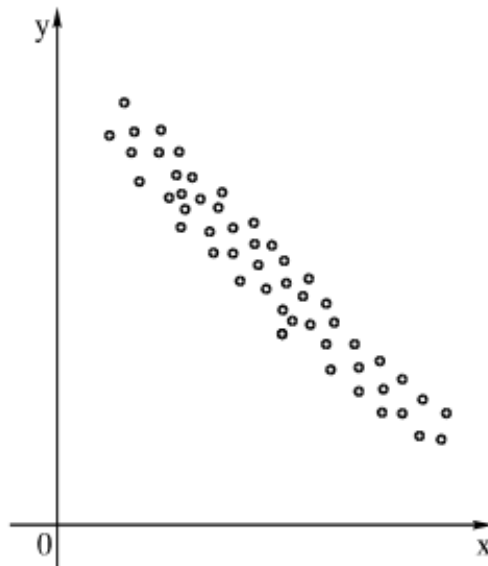
x - nezavisno promenljiva (regresor ili prediktor) (nije slučajna veličina)

Y - zavisna (odzivna ili prom. odgovora) promenljiva (slučajna veličina)

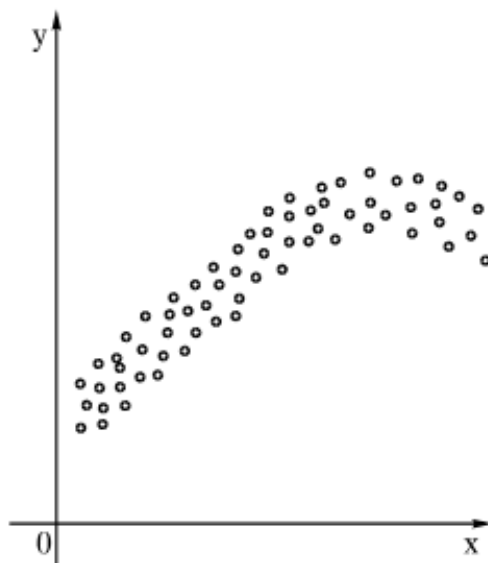
a.) Rastuća linearna zavisnost



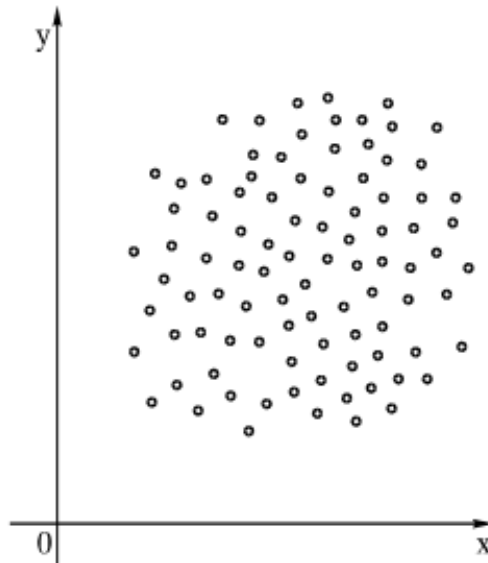
b.) Opadajuća linearna zavisnost



c.) Nelinearna (krivolinijska) zavisnost

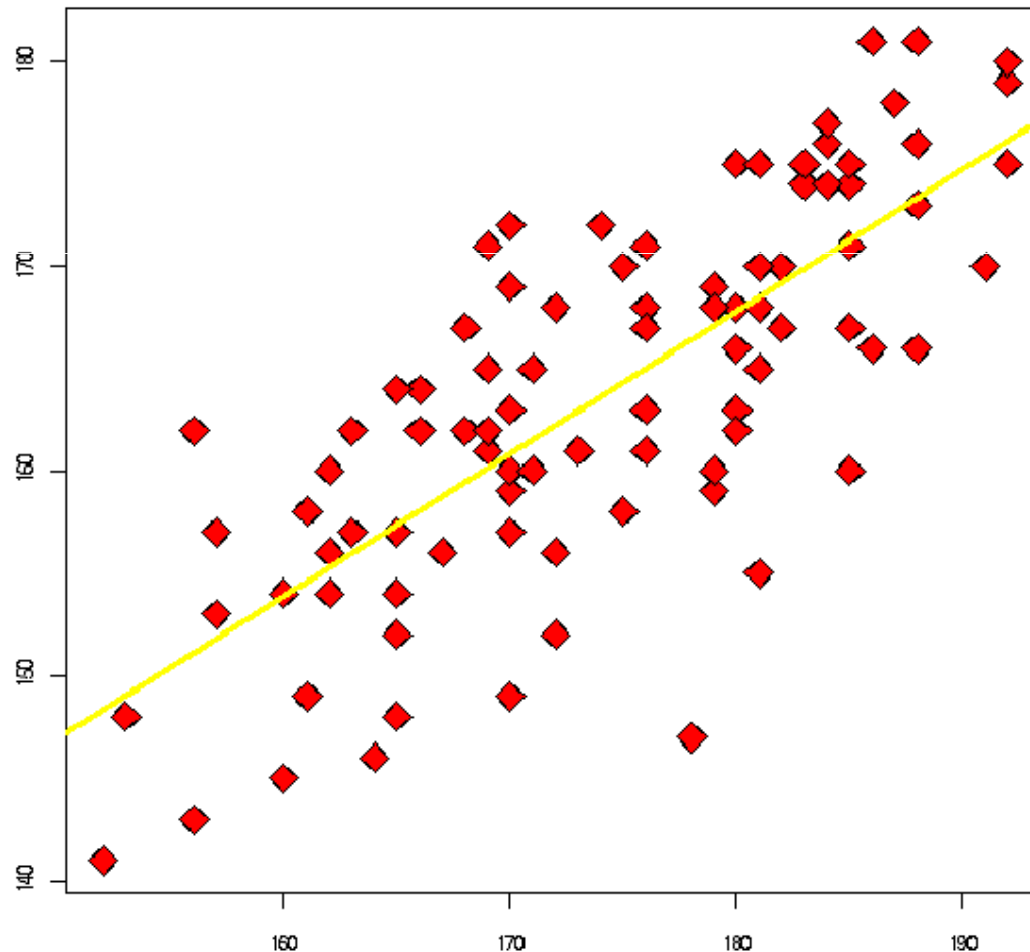


d.) Odsustvo zavisnosti



Da li bi predhodni primer
trebao da bude rastuća ili
opadajuća zavisnost?

Prosta linearna regresija



Kako grafički naći zavisnost između nezavisne promenljive x i zavisne Y (slučajna promenljiva) ?

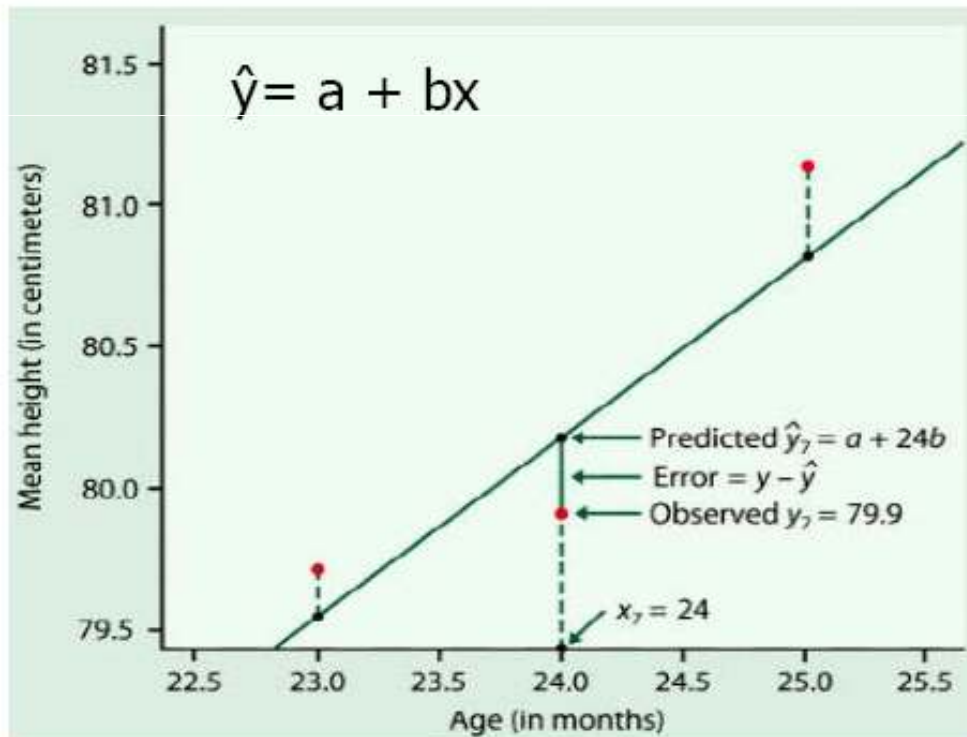
Prava koja na najbolji način aproksimira oblak tačaka se naziva regresiona prava.

Analitički oblik te prave je

$$y = a + bx$$

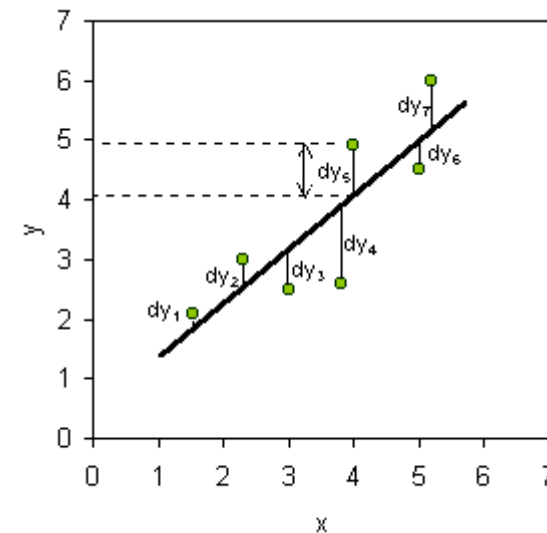
Uočimo razlike između očekivane vrednosti (očitane sa regresione prave) i vrednosti dobijene opažanjem

dy_1, dy_2, \dots, dy_n



Milan Kilibarda

Kakav uslov bi trebalo postaviti da nađemo pravu koja najbolje aproksimira zavisnost između x i Y ?



84

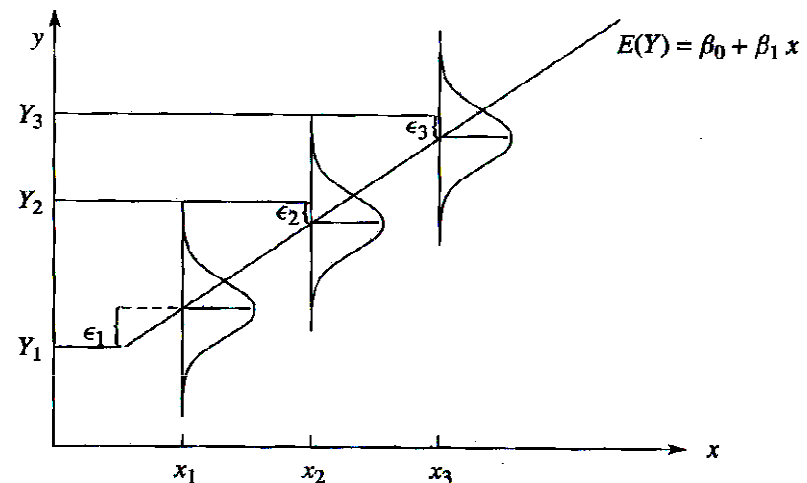
Prosta linearna regresija

$Y = a + bX$ funkcionaln zavisnost

$Y = \beta_0 + \beta_1 x + \varepsilon$ regresiona zavisnost

ε – razlika predikovane i opazane vrednosti

$$\text{sistem}^* \begin{cases} -\varepsilon_1 = \beta_0 + \beta_1 x_1 - Y_1 \\ -\varepsilon_2 = \beta_0 + \beta_1 x_2 - Y_2 \\ \vdots \\ -\varepsilon_n = \beta_0 + \beta_1 x_n - Y_n \end{cases}$$



ε - je slučajna greška sa normalnim rasporedom $N(0, s)$ (značajno za primjenu MNK) ($E(e) = E(-e) = 0$)

$$E[Y] = E[\beta_0 + \beta_1 x + \varepsilon] = \beta_0 + \beta_1 x$$

- Imamo sistem sa n jednačina i 2 tražene veličine.
- U ovom sistemu jednačina su nam nepoznate i slučajne veličine e_i .
- Takav sistem je nedodređen tj. ima beskonačno rešenja.

Prosta linearna regresija

Primer takvog sistema je

$$x+y=0$$

(1,-1; 0,0; 2,-2; itd.)

Dakle, kroz oblak tačaka možemo nacrtati beskonačno pravih, ali mi tražimo onu pravu kod koje je suma kvadrata odstupnaja minimalna (Metod Najmanjih Kvadrata, MNK). Sistem jednačina * uz ovaj uslov ima samo jedno rešenje.

Prosta linearna regresija

Sistem * se može napisati u obliku

$$\begin{cases} v_1 = \beta_0 + \beta_1 x_1 - Y_1 \\ v_2 = \beta_0 + \beta_1 x_2 - Y_2 \\ \vdots \\ v_n = \beta_0 + \beta_1 x_n - Y_n \end{cases} \triangleright \{ \mathbf{v} = \mathbf{A}\mathbf{t} + \mathbf{f}$$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \mathbf{t} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \quad \mathbf{f} = \begin{bmatrix} -Y_1 \\ -Y_2 \\ \vdots \\ -Y_n \end{bmatrix}$$

$\mathbf{v} = \mathbf{A}\mathbf{t} + \mathbf{f}$ – linearni model

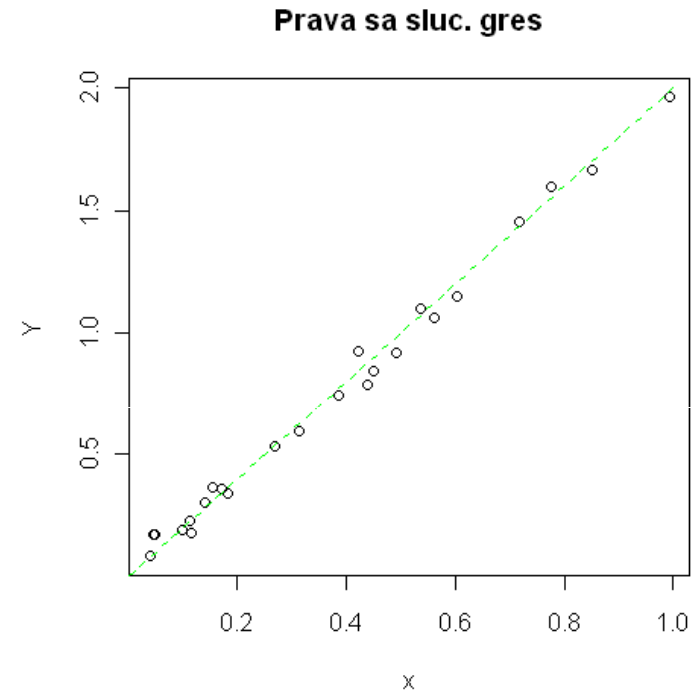
$E(\mathbf{v}) = 0, \mathbf{v} \sim N(0, \sigma)$ – stohasticki model

$$E[Y] = E[\beta_0 + \beta_1 x + \varepsilon] = \beta_0 + \beta_1 x$$

Primer u R-u

```
set.seed(520)
len <- 24
x <- runif(len)
Y <- 2*x + rnorm(len, 0, 0.06)
ds <- data.frame(x = x, Y= Y)
str(ds)
'data.frame': 24 obs. of 2 variables:
 $ x: num 0.1411 0.4925 0.0992 0.0469 0.1131 ...
 $ y: num 0.02586 0.05546 -0.0048 0.0805 0.00764 ...
```

```
plot(Y ~ x, main = "Prava sa sluc. gres")
s <- seq(0, 1, length = 100)
lines(s, 2*s, lty = 2, col = "green")
```



lm - linearni model

```
m <- lm(Y ~ x)
```

```
m
```

Call:

```
lm(formula = Y ~ x)
```

Coefficients:

(Intercept)	x
0.01594	1.94808

Dijagnostifikacija regresionog modela

<code>anova(m)</code>	ANOVA tabela
<code>coefficients(m) ; coef(m)</code>	Koeficijenti modela
<code>confint(m)</code>	Intervali poverenja za koef.
<code>deviance(m)</code>	Suma kvadrata reziduala
<code>effects(m)</code>	Vektor ortogonalnog efekta
<code>fitted(m)</code>	Vektor predikovane Y vrednosti
<code>residuals(m); resid(m)</code>	Reziduali
<code>summary(m)</code>	R ² ; F statistika; standardna greška reziduala (σ) ...
<code>vcov(m)</code>	Varijans-kovarijans matrica parametara

Rezultati modela

coef(m)

```
(Intercept)      x  
0.0159367  1.9480782
```

confint(m)

```
          2.5 %    97.5 %  
(Intercept) -0.01790562 0.04977901  
x           1.87441888 2.02173744
```

Rezultati modela

fitted(m)

1	2	3	4	...
0.29078230	0.97530848	0.20919849	0.10726425	...

resid(m)

1	2	3	4	...
0.014439376	-0.054348301	-0.016565504	0.066898725	...

Rezultati modela

summary(m)

Call:

lm(formula = Y ~ x)

Residuals:

Min	1Q	Median	3Q	Max
-0.083317	-0.032944	-0.006018	0.040948	0.084004

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.01594	0.01632	0.977	0.339
x	1.94808	0.03552	54.848	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04706 on 22 degrees of freedom

Multiple R-squared: 0.9927, Adjusted R-squared: 0.9924

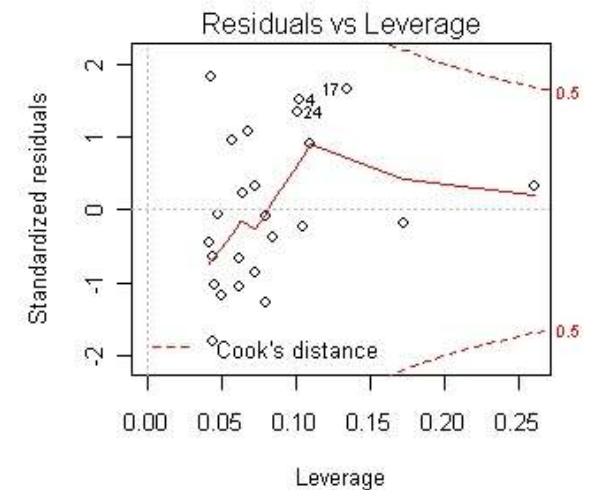
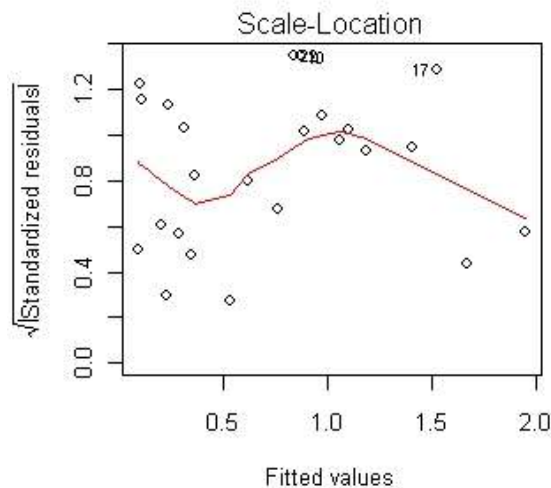
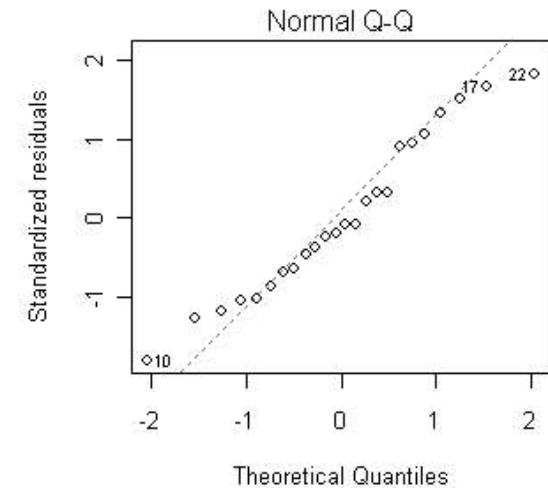
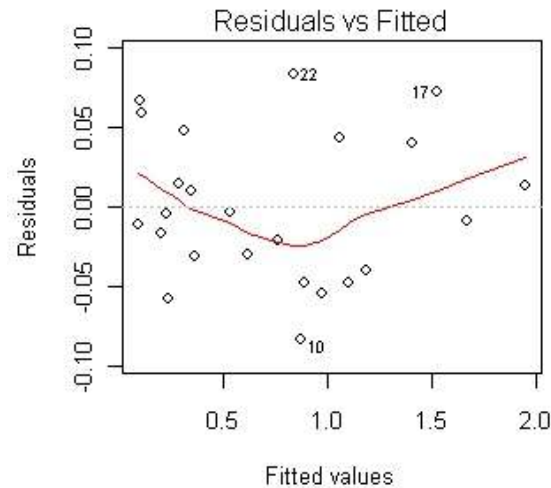
F-statistic: 3008 on 1 and 22 DF, p-value: < 2.2e-16

vcov(m)

	(Intercept)	x
(Intercept)	0.0002662905	-0.0004685335
x	-0.0004685335	0.0012615097

Rezultati modela

```
par(mfrow = c(2, 2))  
plot(m)  
par(mfrow = c(1, 1))
```



Interpretacija

1. Pregled svih reziduala; 10, 17, 22 su moguće grube greške.
2. Q-Q plot - papir verovatnoće. Poredjenje teoretskih kvantila i eksperimentalnih. Pokazuje koliko reziduali slede normalnu raspodelu.
3. Pokazuje centriranost raspodele reziduala.
4. Cook-ove distance pokazuju uticaj pojedinačnih opservacija na model.

Zadavanje modela

$y \sim x$ - Prosta regresija

$y \sim 1 + x$ - Naglašen y-odsečak INTERCEPT

$y \sim -1 + x$ - Regresija kroz koor. početak

$y \sim x1 + x2 + x3$ - Višestruka regresija

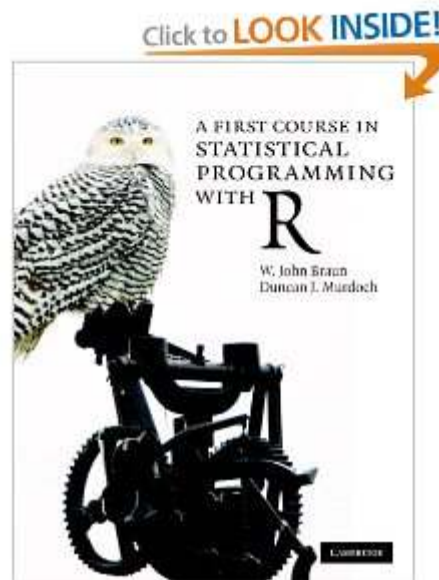
$y \sim x + x^2$ - nelinearna zavisnost (nlm)

Statističke funkcije

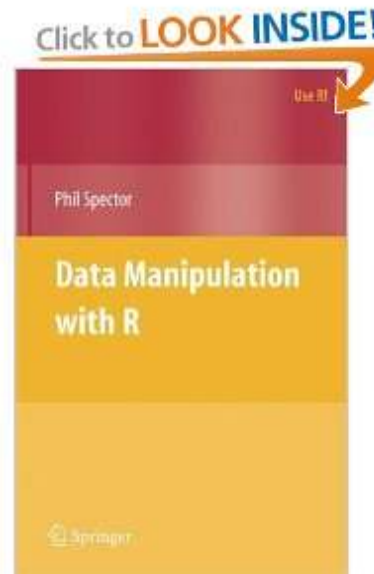
<code>rnorm, dnorm, pnorm, qnorm</code>	Uzorak iz normalne raspodele, gustina, cdf i kvantili
<code>lm, glm, anova</code>	Fitovanje modela
<code>loess, lowess</code>	Fitovanje krive
<code>sample</code>	Resampling (bootstrap, permutation) Uzorkovanje iz skupa.
<code>.Random.seed</code>	Generisanje slučajnih brojeva
<code>mean, median</code>	Mere lokacije
<code>var, cor, cov, mad, range</code>	Deskriptivna stat.
<code>svd, qr, chol, eigen</code>	Linearna algebra

Literatura

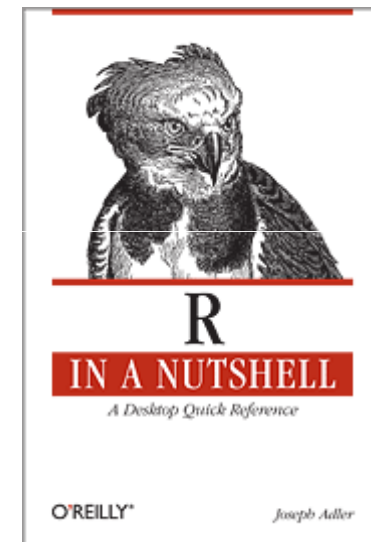
Knjige



**A First Course in Statistical
Programming with R**
W. John Braun, Duncan J.
Murdoch



**Data Manipulation with R (Use
R)**
Phil Spector



**R in a Nutshell
A Desktop Quick
Reference**
Joseph Adler

Tutoriali u pdf-u

- P. Kuhnert & B. Venables, [An Introduction to R: Software for Statistical Modeling & Computing](http://cran.r-project.org/doc/contrib/Kuhnert+Venables-R_Course_Notes.zip)
http://cran.r-project.org/doc/contrib/Kuhnert+Venables-R_Course_Notes.zip
- J.H. Maindonald, [Using R for Data Analysis and Graphics](http://cran.r-project.org/doc/contrib/usingR.pdf)
<http://cran.r-project.org/doc/contrib/usingR.pdf>
- W.J. Owen, [The R Guide](http://www.mathcs.richmond.edu/~wowen/TheRGuide.pdf)
<http://www.mathcs.richmond.edu/~wowen/TheRGuide.pdf>
- D. Rossiter, [Introduction to the R Project for Statistical Computing for Use at the ITC](http://cran.r-project.org/doc/contrib/Rossiter-RIntro-ITC.pdf)
<http://cran.r-project.org/doc/contrib/Rossiter-RIntro-ITC.pdf>
- W.N. Venables & D. M. Smith, [An Introduction to R](http://cran.r-project.org/doc/manuals/R-intro.pdf)
<http://cran.r-project.org/doc/manuals/R-intro.pdf>

Web links

- Paul Geissler's excellent R tutorial

<http://www.fort.usgs.gov/BRDScience/LearnR.htm>

- Dave Robert's Excellent Labs Ekološke
analize

<http://ecology.msu.montana.edu/labds/R/labs/>

- Excellent Tutorials by David Rossiter

<http://www.itc.nl/personal/rossiter/pubs/list.html>

- Excellent tutorial on nearly every aspect
of R

<http://www.statmethods.net/index.html>

- Introduction to R by Vincent Zoonekynd

http://zoonek2.free.fr/UNIX/48_R/all.html

Web links

- <http://casoilresource.lawr.ucdavis.edu/drupal/node/100>
- [R Concepts and Data Types](http://www.stat.wisc.edu/~deepayan/SIBS2005/slides/language-overview-4.pdf)
presentation by Deepayan Sarkar
- [Interpreting Output From lm\(\)](http://www.rni.helsinki.fi/~pek/s-tools/lm-more.r)
- [The R Wiki](http://cran.r-project.org/doc/manuals/R-data.html)
- [Import / Export Manual](http://cran.r-project.org/doc/contrib/Short-refcard.pdf)
- [R Reference Cards](http://cran.r-project.org/doc/contrib/Short-refcard.pdf)

Web links

- KickStart <http://cran.r-project.org/doc/contrib/Lemon-kickstart/index.html>
- Hints on plotting data in R
<http://www.stat.auckland.ac.nz/~paul/RGraphics/rgraphics.html>
- Regression and ANOVA
cran.r-project.org/doc/contrib/Faraway-PRA.pdf
- A Handbook of Statistical Analyses Using R (Brian S. Everitt and Torsten Hothorn)
<http://cran.r-project.org/web/packages/HSAUR>