

# Statistika

IIS SIIT

FTN Novi Sad

## 1 Instalacija R i R-studio (za MS Windows)

### 1.1 R

Na stranici <https://cran.r-project.org/bin/windows/base/> skinite poslednju verziju R-a. Trenutno je to <https://cran.r-project.org/bin/windows/base/R-3.5.1-win.exe> Pokrenite skinutu aplikaciju i pratite uputstva. Pokretanje instaliranog programa je intuitivno jasno.

### 1.2 R-studio

Na stranici <https://www.rstudio.com/products/rstudio/download/#download> skinite poslednju verziju R-studija. Trenutno je to <https://download1.rstudio.org/RStudio-1.1.456.exe> Pokrenite skinutu aplikaciju i pratite uputstva. R-studio je komforno okruženje za korišćenje R-a.

### 1.3 IOS, Linux

Slično kao za Windows.

### 1.4 Kratak kurs R-a

Odličan uvod za korišćenje R-a ima u:

<https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>

### 1.5 Skidanje podataka za vežbanje

<https://sites.google.com/site/radojkaciganovicftn/fajlovi/Data.zip?attredirects=0&d=1>

## 2 Osnovne operacije u R-u

### 2.1 Pokretanje R-a (MS Windows)

1. Napraviti direktorijum u kome ćete smeštati R fajlove
2. Podesiti Properties od ikone R-studio-a: Start In podesiti na punu putanju do napravljenog direktorijuma
3. Pokrenuti R-studio dvoklikom na podešenu ikonu
4. Isprobati u komandnom promptu

```
x<-sin(1)
q()
```

Save Workspace Image: Yes (biće sačuvan u fajlu .RData, istorija komandi u .Rhistory)

5. Ponovo pokrenuti R

```
objects ()
x
```

treba da bude  $x = 0.841471$ .

### 2.2 Praktičan unos podataka - CSV fajlovi (spreadsheet programi)

Skinuti data set knjige:

Mathematical Statistics with Resampling and R, Chihara, Hersterberg.

Raspakovati fajl `FlightDelays.csv` u folder odakle se pokreće R.

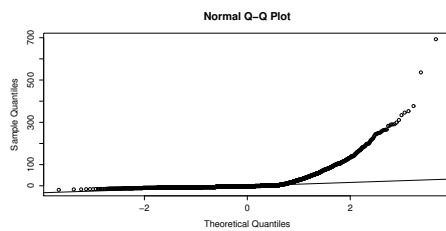
ID	Carrier	FlightNo	Destination	DepartTime	Day	Month	FlightLength	Delayed	Delayed30
1	UA	403	DEN	4-8am	Fri	May	281	-1	No
2	UA	45	DEN	8-Noon	Fri	May	277	102	Yes
...									

### 2.3 Manipulisanje podataka

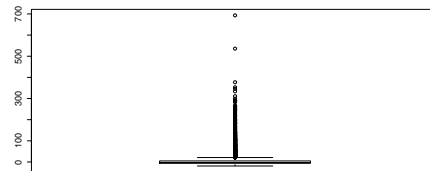
```
setwd("C:/Users/zoran/Documents/R") # Postavljanje radnog direktorijuma
rm(list=ls ()) # Brisanje svih workspace promenljivih
load(".RData") # Ucitavanje promenljivih iz R formata
FlightDelays <- read.csv("FlightDelays.csv") # Ucitavanje CSV tabele
save.image("nov.RData") # Cuvanje Work space u fajlu nov.RData
save(FlightDelays, file="Fl.rda") # Cuvanje FlightDelays u fajlu Fl.rda
summary(FlightDelays) # Osnovna deskriptivna statistika
edit(FlightDelays) -> FlightDelays # Otvaranje data editora
str(FlightDelays) # Daje strukturu podataka
names(FlightDelays) # Izlistava promenljive u podacima
head(FlightDelays) # Prvih 6 redova tabele podataka
tail(FlightDelays) # Poslednjih 6 redova
FlightDelays [1:10,1:3] # Prvih 10 redova, prve tri promenljive
```

## 2.4 Osnovne R komande

```
# Prebacivanje kolone u niz
delay <- FlightDelays$Delay
# Duzina niza
n<-length(delay)
# Nalazenje kvantila normalne N(0,1)
# raspodele za redom elemente niza
qdelay<-qnorm((1:n)/n)
# qqplot za normalnu N(0,1) raspodelu
plot(qdelay, sort(delay))
# Isto, samo sa jednom komandom
qqnorm(delay)
# Povlacenje prave linije koja bi
# odgovarala datim kvantilima
qqline(delay)
# Vidimo da delay nema Normalnu raspodelu
summary(delay)
# Interkvantilni razmak za delay
IQR(delay)
```



```
boxplot(delay)
```



```
boxplot(Delay ~ Day, data = FlightDelays)
```

