



**NAPOMENA: Ovo je radna (početna) verzija materijala ... pa su greške moguće !!!**

# Master Industry 4.0

## UVOD U NAUKU O PODACIMA

### R Elements

### BLOK 2f



- o Dokument je ISKLJUČIVO namenjen za internu upotrebu za studente Master akademskih studija informatike na Prirodno-matematičkom fakultetu u Kragujevcu za predmet Uvod u nauku o podacima.
- o Dokument sadrži i odabrane segmente iz navedene literature sa okvirnim prevodom.
- o Reč je o prvoj verziji dokumenta pa su moguće greške.
- o Za detaljnije proučavanje prikazane materije studentima se preporučuje originalna navedena literatura.
- o Za referenciranje i citiranje OBAVEZNO je navođenje originalne literature.

## Literatura

1. Adler, J. R in a nutshell: A desktop quick reference. " O'Reilly Media, Inc.", 2010
2. Pathak, Manas A., Beginning Data Science with R. Springer, 2014.
3. Schutt, Rachel, and Cathy O'Neil. Doing data science: Straight talk from the frontline. " O'Reilly Media, Inc.", 2013.
4. Wickham, Hadley, and Garrett Grolemund, R za statističku obradu podataka, Mikro knjiga, 2017.

## 1. Dobijanje pomoći

R uključuje i sistem za pomoć oko informacija o instaliranim paketima. Za dobijanje pomoći za funkcije, na primer, `glm` trebalo bi da ukucate:

```
> help(glm)
```

Ili ekvivalentno:

```
> ?glm
```

Za pomoć u vezi operatora, stavite operator potrebno je da stavite operator pod jednostruke navodnike. Na primer:

```
> ?`+`
```

Primere za određenu funkciju možete dobiti preko funkcije `example`:

```
> example(glm)
```

Moguća je i pretraga help sistema pomoću ključnih reči:

```
> help.search("regression")
```

Ili preko prečice:

```
> ??regression
```

Za pomoć za određeni paket kucajte, na primer za paket `grDevices`:

```
> library(help="grDevices")
```

Neki paketi poseduju makar jednu vinjetu. Vinjeta je kratak document koji opisuje kako koristiti paket, zajedno sa primerima korišćenja. Recimo za dobijanje vinjete za paket `dplyr` kucajte:

```
> vignette("dplyr")
```

Da biste izlistali vinjete za sve instalirane ili sve prikazane pakete kucajte:

```
> vignette(all=FALSE) #lista vinjeta samo za prikazane pakete  
> vignette(all=TRUE) #lista vinjeta za sve instalirane pakete
```

## 2. R paketi

Paket predstavlja skup srodnih funkcija, zajedno sa pomoćnim datotekama, i datotekama sa podacima koje dolaze zajedno sa paketom. Paketi u R su slični modulima u Perlu, bibliotekama u C/C++, i klasama u Javi.

Tipično sve funkcije u paketu su povezane: na primer paket `stats` sadrži funkcije za statističke analize. Da bi koristili paket potrebno je da ga učitate u R.

R nudi ogroman broj paketa, dizajniranih za širok spektar sektora i primena: za grafiku, za obradu statističkih testova, pakete sa najnovijim metodama mašinskog učenja...

Neki od paketa su uključeni u R: samo treba da R-u stavite do znanja da želite da ih koristite. Većina ostalih paketa su dostupni iz javnih repozitorijuma paketa. Možete kreirati čak i svoje sopstvene pakete.

### 2.1. Listing paketa u lokalnim bibliotekama

Da bi dobili listu paketa koji su podrazumevano učtani, koristi se komanda `getOption`:

```
> getOption("defaultPackages")  
[1] "datasets"      "utils"         "grDevices"
```

```
[4] "graphics"      "stats"      "methods"  
[7] "rpart"         "lattice"    "RevoScaleR"  
[10] "mrsdeploy"     "MicrosoftML" "RevoMods"  
[13] "RevoUtils"     "RevoUtilsMath"
```

Ova komanda izostavlja base paket, koji je uvek učitан. Ukoliko želite da vidite listu trenutno učitаних paketa, možete se koristiti komandom `(.packages())` (obratite pažnju na spoljne zagrade!):

```
> (.packages())  
[1] "ggplot2"                "nutshell"  
[3] "nutshell.audioscrobbler" "nutshell.bbdb"  
[5] "RevoUtilsMath"         "RevoUtils"  
[7] "RevoMods"              "MicrosoftML"  
[9] "mrsdeploy"             "RevoScaleR"  
[11] "lattice"               "rpart"  
[13] "stats"                 "graphics"  
[15] "grDevices"            "utils"  
[17] "datasets"              "methods"  
[19] "base"
```

Da biste videli sve dostupne pakete, koristite opciju `all.available=TRUE`, zajedno sa `(.packages())`:

```
> (.packages(all.available=TRUE))
```

Takođe možete u istu svrhu pokrenuti `library()` funkciju bez argumenata:

```
> library()
```

Evo liste paketa uključenih u standardnu CRAN R instalaciju:



Package name	Loaded by default	Description
base	✓	Basic functions of the R language, including arithmetic, I/O, programming support
boot		Bootstrap resampling
class		Classification algorithms, including nearest neighbors, self-organizing maps, and learning vector quantization
cluster		Clustering algorithms
codetools		Tools for analyzing R code
compiler		Byte code compiler for R
datasets	✓	Some famous data sets
foreign		Tools for reading data from other formats, including Stata, SAS, and SPSS files
graphics	✓	Functions for base graphics
grDevices	✓	Device support for base and grid graphics, including system-specific functions
grid		Tools for building more sophisticated graphics than the base graphics
KernSmooth		Functions for kernel smoothing
lattice		An implementation of Trellis graphics for R: prettier graphics than the default graphics
MASS		Functions and data used in the book <i>Modern Applied Statistics with S</i> by Venables and Ripley; contains a lot of useful statistics functions
methods	✓	Implementation of formal methods and classes introduced in S version 4 (called S4 methods and classes)
mgcv		Functions for generalized additive modeling and generalized additive mixed modeling
nlme		Linear and nonlinear mixed-effects models
nnet		Feed-forward neural networks and multinomial log linear models
parallel		Support for parallel computation, including random-number generation.
rpart		Tools for building recursive partitioning and regression
spatial		Functions for Kriging and point pattern analysis
splines		Regression spline functions and classes
stats	✓	Functions for statistics calculations and random number generation; includes many common statistical tests, probability distributions, and modeling tools
stats4		Statistics functions as S4 methods and classes



survival		Survival analysis functions
tcltk		Interface to Tcl/Tk; used to create platform-independent UI tools
tools		Tools for developing packages
utils	✓	A variety of utility functions for R, including package management, file reading and writing, and editing

Slika 2.1.1. Podrazumevani R paketi

## 2.2. Instalacija i učitavanje novih paketa

Kao primer, instalirajmo i aktivirajmo `tidyverse` paket. Pri instalaciji će se takođe instalirati paketi koji su preduslov da paket bude operativan:

```
> install.packages("tidyverse")
also installing the dependencies 'mnormt', 'rematch', 'psych', 'openssl', 'cellranger', 'selectr', 'tidyselect', 'broom', 'forcats', 'haven', 'httr', 'hms', 'lubridate', 'modelr', 'purrr', 'readr', 'readxl', 'rvest', 'tidyr', 'xml2'
```

Moguće je i eksplicitno definisati direktorijum u koji će se paket instalirati, i/ili iz koga će paket biti učitani, što je korisna opcija ukoliko, recimo nemate Administratorski nalog na nekom računaru.

```
> install.packages("zoo", lib="I:/MojiRPaketi/") #instalacija paketa
> library("zoo", lib.loc="I:/MojiRPaketi/") #učitavanje/aktiviranje paketa
```

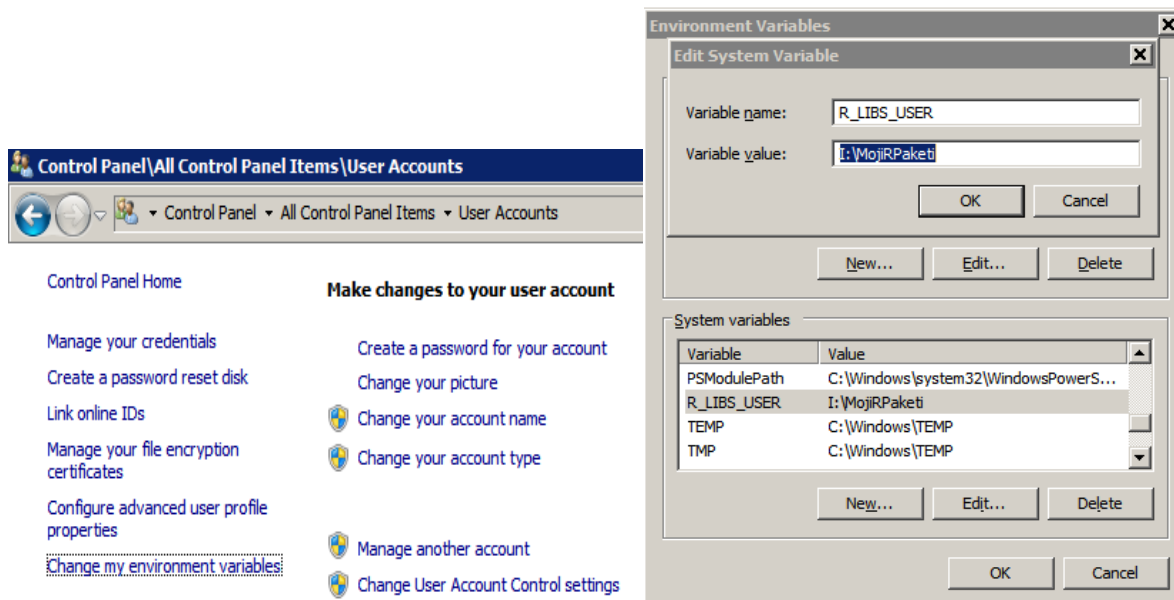
Ovde postoji potencijalna smetnja. Pri učitavanju ovako instaliranog paketa uvek je neophodno obezbediti i `lib.loc` argument. Da bismo ovo izbegli upotrebimo sledeće rešenje:

```
#ovim smo dodajemo još jednu lokaciju gde će R tražiti pakete koji mogu biti učitani
> .libPaths(c(.libPaths(),"I:/MojiRPaketi"))
#proverimo koji su to direktorijumi gde će R tražiti pakete koji mogu biti učitani
> .libPaths()
[1] "I:/RKM/R_SERVER/library" "I:/MojiRPaketi"
```

Slično se može postići (na Windowsu) sledećom procedurom:

- Kreirajte direktorijum u kojem želite da skladištite R pakete, recimo I:\MojiRPaketi
- Kliknite Start --> Control Panel --> User Accounts --> Change my environment variables.
- Ukoliko vidite stavku `R_LIBS_USER`, obeležite je i kliknite *Edit*. U suprotnom, kliknite *New*. U polje *Variable* unesite `R_LIBS_USER` (ukoliko veće nije tu). U polje *Value* unesite putanju do željenog direktorijuma, I:\MojiRPaketi. Moguće je definisati više putanja odvojenih sa `;`.

Procedura je prikazana I na slici.



Slika 2.2.1. Environment variables dijalog sa potrebnim izmenama u cilju promene podrazumevanih putanja za instalaciju i učitavanje R paketa

Neki paketi ne mogu se preuzeti sa CRAN repozitorijuma. Recimo Twitterov paket `anomalous`, ili Netflixov paket `RAD` – robust anomaly detection. U ovakvim slučajevima koristimo devtools biblioteku koja sadrži alate za instalaciju paketa sa repozitorijuma kao što je github.

```
> install.packages("devtools") # sadrži potrebne alate za instalaciju paketa sa githuba
> devtools::install_github("twitter/AnomalyDetection") #instalacija paketa AnomalyDetection
> library(AnomalyDetection) #učitavanje paketa AnomalyDetection
> library(withr) #sadrži alate neophodne za instalaciju paketa sa githuba u specifikovani direk
torijum
```

```
#Donju komandu je neophodno pokrenuti dva puta ukoliko R build tools nije instaliran
> withr::with_libpaths(new = "I:/MojIRPaketi", devtools::install_github(repo = "Surus", usernam
e = "Netflix", subdir = "resources/R/RAD"))
```

```
> library(RAD) #učitavanje paketa RAD
```

Da bi „otkačili” učitanu paket, recimo paket `zoo`, koristimo `detach` funkciju:

```
> detach("package:zoo", unload=TRUE)
Warning message:
'zoo' namespace cannot be unloaded:
 namespace 'zoo' is imported by 'quantmod', 'TTR', 'xts', 'tseries' so cannot be unloaded
```

Ukoliko je argument `unload` postavljen na TRUE, nakon odvajanja paketa, biće pokušano rasterećenje prostora imena. U ovom slučaju paket je odvojen, ali rasterećenje prostora imena je neuspelo. Listu prikačenih paketa (i R objekata kao što su data frame-ovi) možemo dobiti i pomoću `search()` funkcije. U tom slučaju možemo koristiti index paketa da bi ga otkačili:

```
> search()
[1] ".GlobalEnv" "package:AnomalyDetection"
[3] "package:withr" "package:ggplot2"
```



```
[5] "package:nutshell"           "package:nutshell.audioscrobbler"  
[7] "package:nutshell.bbdb"      "tools:rstudio"  
[9] "package:RevoUtilsMath"     "package:RevoUtils"  
[11] "package:RevoMods"          "package:MicrosoftML"  
[13] "package:mrsdeploy"         "package:RevoScaler"  
[15] "package:lattice"           "package:rpart"  
[17] "package:stats"             "package:graphics"  
[19] "package:grDevices"         "package:utils"  
[21] "package:datasets"          "package:methods"  
[23] "Autoloads"                 "package:base"
```

```
> detach(4, unload = TRUE) #otkačuje paket ggplot2
```

### 2.3. Уklanjanje paketa

Za uklanjanje paketa koristimo funkciju `remove.packages()`.

```
> remove.packages("ggplot2") #nije specificovana putanja do paketa. Koristiće se podrazumevana  
> remove.packages("zoo", lib="I:/MojiRPaketi") #specificovan direktorijum gde se paket nalazi
```

## 3. Unos, učitavanje, izmena, i skladištenje podataka. EDA. ggplot.

Ako unosite mali broj podataka, unošenje podataka direktno preko R editora, može biti dobar pristup. Postoje dva različita načina za direktni unos podataka.

### 3.1. Unos podataka korišćenjem R komandi / elementarni primeri

Kao što smo ranije videli, vektore možemo kreirati sa `c()` funkcijom:

```
> salary <- c(18700000, 14626720, 14137500, 13980000, 12916666)  
> position <- c("QB", "QB", "DE", "QB", "QB")  
> team <- c("Colts", "Patriots", "Panthers", "Bengals", "Giants")  
> name.last <- c("Manning", "Brady", "Pepper", "Palmer", "Manning")  
> name.first <- c("Peyton", "Tom", "Julius", "Carson", "Eli")
```

Često je pogodnije držati ove vektore zajedno u data frame-u. Za kreiranje data frame-a, koristite funkciju `data.frame` radi spajanja vektora:

```
> top.5.salaries <- data.frame(name.last, name.first, team, position, salary)  
> top.5.salaries  
  name.last name.first   team position  salary  
1  Manning   Peyton   Colts      QB 18700000  
2   Brady     Tom   Patriots      QB 14626720  
3   Pepper   Julius   Panthers      DE 14137500  
4   Palmer   Carson   Bengals      QB 13980000  
5  Manning     Eli    Giants      QB 12916666
```

### 3.2. Uvoz podataka iz eksternih datoteka. EDA primer. /// PRIMER za navikavanje na R okruženje!!!

Radićemo sa datasetom imenovanim *nyt1*, dostupnim na adresi <http://stat.columbia.edu/~rachel/datasets/nyt1.csv> ili [https://github.com/oreillymedia/doing\\_data\\_science/blob/master/dds\\_datasets.zip](https://github.com/oreillymedia/doing_data_science/blob/master/dds_datasets.zip)

Ovaj dataset sadrži informacije o prikazanim reklamama i zabeleženim klikovima na web stranici *New York Times* novina, za jedan dan u maju 2012. godine. Svaki red u datasetu predstavlja jednog korisnika. Dataset ima 5 kolona: godine, pol (0=Zenski, 1=Muski), broj utisaka, broj klikova, ulogovan korisnik(1=ulogovan,0=nije ulogovan). Pod utiskom ili impresijom ovde se misli na **dogadjaj** preuzimanja reklamnog sadržaja sa njegovog izvora, radi učitavanja na web stranicu kod korisnika (reklamni banner ili nešto drugo). Zadaci su sledeći:

- ✓ Učitati dataset sa njegovog izvora u data frame `data1`.
- ✓ Kreirati novu kategoričku promenljivu, `katgod`, koja će kategorisati korisnike prema starosti u sledeće kategorije: „(-Inf,0]”, „(0,18]”, „(18,24]”, „(24,34]”, „(34,44]”, „(44,54]”, „(54,64]”, „(64, Inf]”.
- ✓ Grafički prikazati raspodelu broja impresija i Click-Through Rate (CTR= #klikova/#impresija) za ovih osam kategorija.
- ✓ Definisati novu promenljivu radi segmentacije ili kategorizacije korisnika prema njihovom ponašanju u klikovanju na reklame.
- ✓ Istražiti podatke i napraviti vizuelna i kvantitativna poređenja za određene segmente/demografiju korisnika (recimo ženki korisnici mlađji od 18 godina prema muškim korisnicima mlađim od 18 godina, ili ulogovani prema neulogovanim korisnicima).
- ✓ Kreirati metrike/merjenja/statističke pokazatelje koji će sumarno predstaviti podatke. Recimo CTR, kvantili, aritmetička sredina, medijana, varijansa, maksimum (za sve korisnike, ili po segmentima korisnika).

```
# CTRL-L briše shell
> rm(list=ls()) #oslobodiće RAM uklanjanjem svih objekata iz tekućeg radnog prostora
> getwd() #prikazuje tekući direktorijum
[1] "C:/Users/srdjan/Documents"
> setwd("e:\\dskg") # Promeni tekući direktorijum na zadati
> setwd("e:/dskg") # Ekvivalentna komanda
> setwd("e:dskg") # Neispravno jer je \ escape znak
> # Preuzmi i učitaj u data frame: data1 nyt1.csv sa datog urla
> data1 <- read.csv(url("http://stat.columbia.edu/~rachel/datasets/nyt1.csv"))
> write.csv(data1,file = "proba1.csv",row.names = F) #Snimi kao csv u setovani direktorijum
> data1 <- read.csv(file="proba1",header = T,sep = ",") #Učitavanje iz lokalnog fajla
> head(data1,n=5) # prvih 5 zapisa
```

```
Age Gender Impressions Clicks Signed_In
1 36 0 3 0 1
2 73 1 3 0 1
3 30 0 3 0 1
4 49 1 3 0 1
5 47 1 11 0 1
> tail(data1,n=5) # poslednjih 5 zapisa
```

```
Age Gender Impressions Clicks Signed_In
458437 0 0 2 0 0
458438 0 0 4 0 0
```

```
458439 72 1 5 0 1
458440 0 0 5 0 0
458441 0 0 3 0 0

> names(data1)
[1] "Age" "Gender" "Impressions" "Clicks" "Signed_In"

> names(data1) <- c("Godine", "Pol", "Utisci", "Klikovi", "KorisnikUlogovan")

> names(data1)
[1] "Godine" "Pol" "Utisci" "Klikovi" "KorisnikUlogovan"
> names(data1)[3] <- "Impresije"
> head(data1)

  Godine Pol Impresije Klikovi Ulogovan
1     36  0         3         0         1
2     73  1         3         0         1
3     30  0         3         0         1
4     49  1         3         0         1
5     47  1        11         0         1
6     47  0        11         1         1

> data1$katgod <- cut(data1$Godine, c(-Inf, 0, 18, 24, 34, 44, 54, 64, Inf))
> data1$altkatgod <- cut(data1$Godine, breaks=c(-Inf, 0, 18, 24, 34, 44, 54, 64, Inf), labels = c("kat1", "kat2", "kat3", "kat4", "kat5", "kat6", "kat7", "kat8"))
> head(data1)

  Godine Pol Impresije Klikovi KorisnikUlogovan katgod altkatgod
1     36  0         3         0         1 (34,44] kat5
2     73  1         3         0         1 (64, Inf] kat8
3     30  0         3         0         1 (24,34] kat4
4     49  1         3         0         1 (44,54] kat6
5     47  1        11         0         1 (44,54] kat6
6     47  0        11         1         1 (44,54] kat6

> typeof(data1$Pol)
[1] "integer"
> unique(data1$Pol)
[1] 0 1
> data1$polFakt[data1$Pol==0] <- "Zenski"
> data1$polFakt[data1$Pol==1] <- "Muski"
> data1$polfakt2 <- ifelse(data1$Pol == 0, "Zenski", "Muski")
> head(data1)

  Godine Pol Impresije Klikovi KorisnikUlogovan katgod altkatgod polFakt polfakt2
1     36  0         3         0         1 (34,44] kat5 Zenski Zenski
2     73  1         3         0         1 (64, Inf] kat8 Muski Muski
3     30  0         3         0         1 (24,34] kat4 Zenski Zenski
4     49  1         3         0         1 (44,54] kat6 Muski Muski
```

```
5 47 1 11 0 1 (44,54] kat6 Muski Muski
6 47 0 11 1 1 (44,54] kat6 Zenski Zenski
```

```
> summary(data1)
```

```
      Godine      Pol      Impresije      Klikovi
Min.   : 0.00   Min.   :0.000   Min.   : 0.000   Min.   :0.00000
1st Qu.: 0.00   1st Qu.:0.000   1st Qu.: 3.000   1st Qu.:0.00000

Median : 31.00   Median :0.000   Median : 5.000   Median :0.00000
Mean   : 29.48   Mean   :0.367   Mean   : 5.007   Mean   :0.09259
3rd Qu.: 48.00   3rd Qu.:1.000   3rd Qu.: 6.000   3rd Qu.:0.00000
Max.   :108.00   Max.   :1.000   Max.   :20.000   Max.   :4.00000
```

```
KorisnikUlogovan  katgod      altkatgod      polFakt
Min.   :0.0000   (-Inf,0]:137106   kat1   :137106   Length:458441

1st Qu.:0.0000   (34,44] : 70860   kat5   : 70860   Class :character
Median :1.0000   (44,54] : 64288   kat6   : 64288   Mode  :character

Mean   :0.7009   (24,34] : 58174   kat4   : 58174
3rd Qu.:1.0000   (54,64] : 44738   kat7   : 44738
Max.   :1.0000   (18,24] : 35270   kat3   : 35270
                (Other) : 48005   (Other): 48005
```

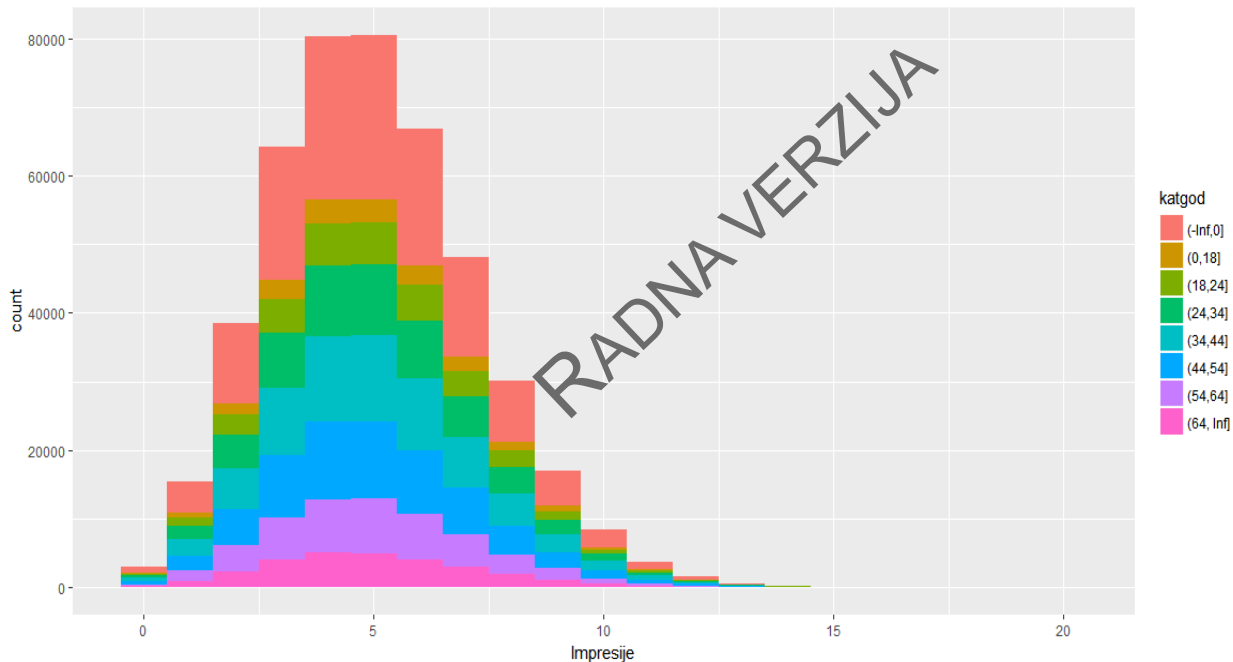
```
polfakt2
Length:458441
Class :character
Mode :character
```

```
> install.packages("doBy")
> library(doBy)
> siterange <- function(x){c(length(x), min(x), mean(x), max(x))}
> summaryBy(Godine~katgod, data=data1, FUN=siterange)
```

```
      katgod Godine.FUN1 Godine.FUN2 Godine.FUN3 Godine.FUN4
1 (-Inf,0]      137106           0      0.00000           0
2 (0,18]        19252           7     16.03350          18
3 (18,24]       35270          19     21.26904          24
4 (24,34]       58174          25     29.50335          34
5 (34,44]       70860          35     39.49468          44
6 (44,54]       64288          45     49.49258          54
7 (54,64]       44738          55     59.49819          64
8 (64, Inf]     28753          65     72.98870         108
```

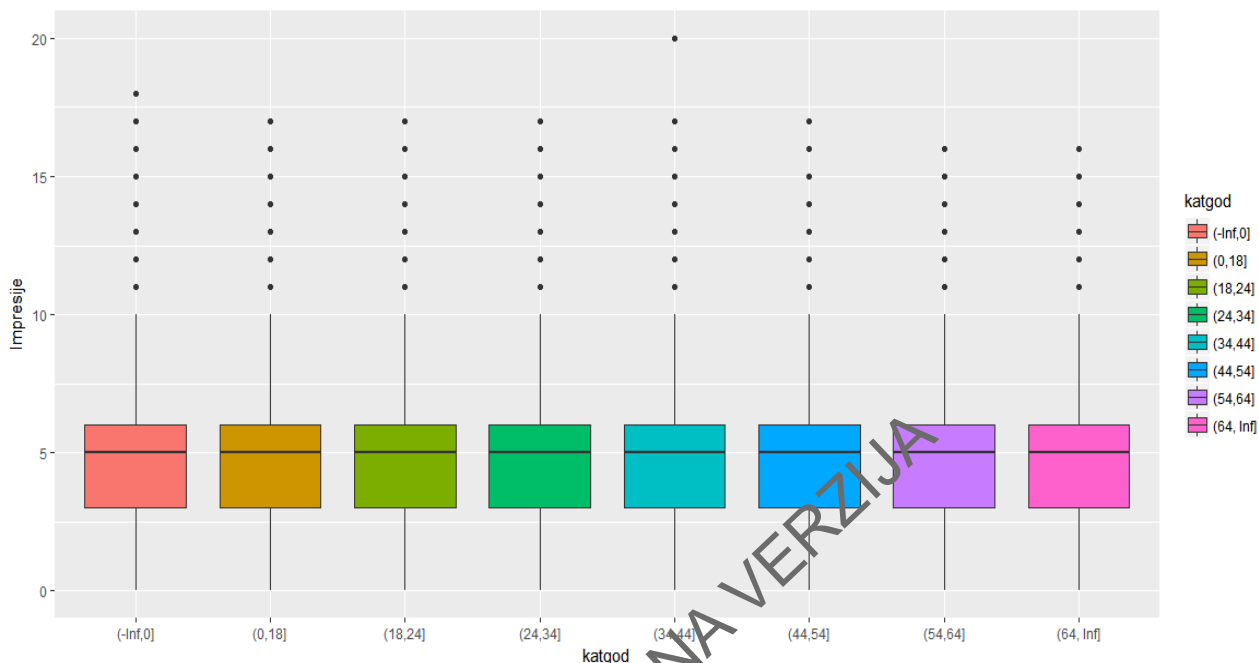
```
> summaryBy(Pol+KorisnikUlogovan+Impresije+Klikovi~katgod,data =data1)
      katgod Pol.mean KorisnikUlogovan.mean Impresije.mean Klikovi.mean
1 (-Inf,0] 0.0000000           0      4.999657 0.14207985
2 (0,18] 0.6421151           1      4.998961 0.13105132
3 (18,24] 0.5338531           1      5.006635 0.04845478
```

```
4 (24,34] 0.5321621 1 4.993829 0.05048647
5 (34,44] 0.5316963 1 5.021507 0.05167937
6 (44,54] 0.5289790 1 5.010406 0.05027377
7 (54,64] 0.5361885 1 5.022308 0.10183736
8 (64, Inf] 0.3632664 1 5.012347 0.15128856
> install.packages("ggplot2")
> library(ggplot2)
> library(tidyverse)
> ggplot(data1, aes(x=Impresije, fill=katgod))+geom_histogram(binwidth=2)
```



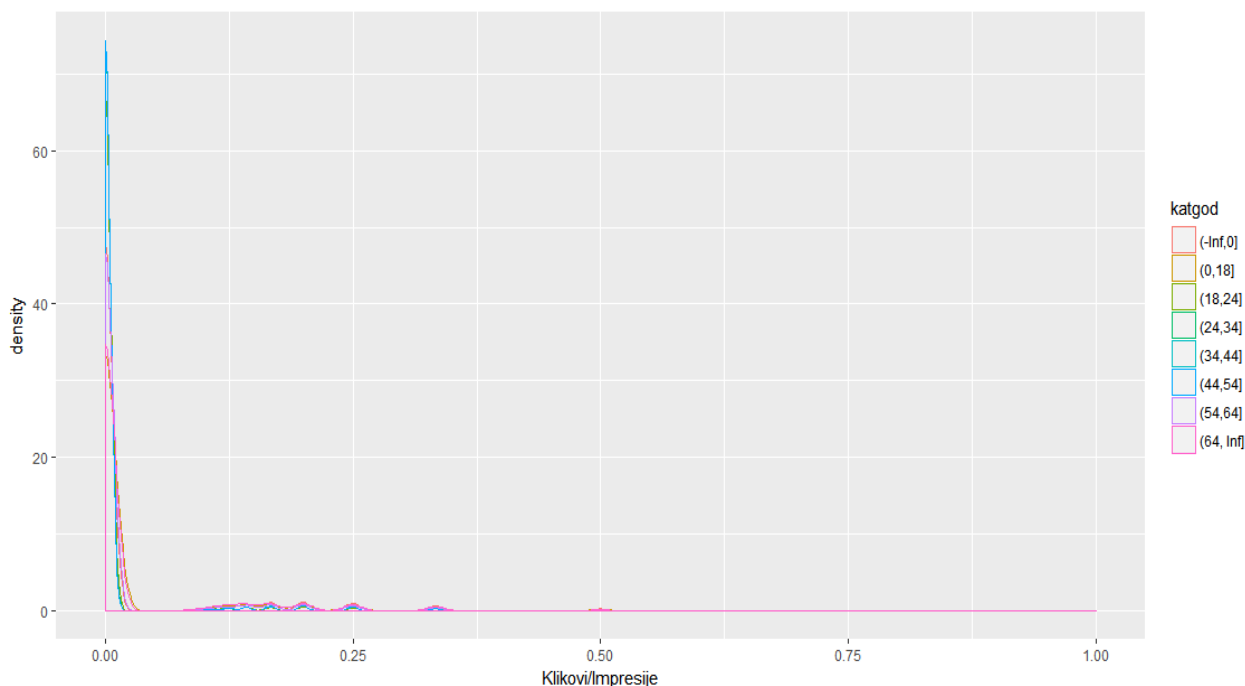
Slika 3.2.1. Histogram utisaka za nyt dataset dan 1

```
> ggplot(data1, aes(x=katgod, y=Impresije, fill=katgod))+geom_boxplot()
```



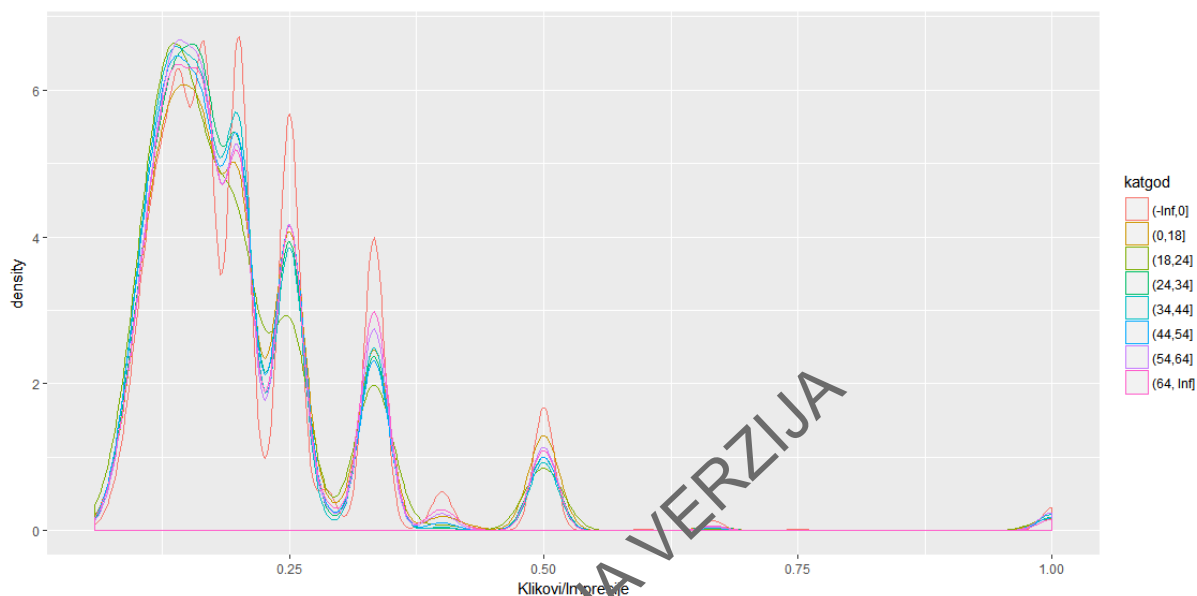
Slika.1.2.2. boxplot utisaka nyt dataset dan 1

```
> data1$imaimpresija <- cut(data1$Impresije,c(-Inf,0,Inf))
> write.csv(x=data1,file="backup7.csv",row.names = F)
> summaryBy(Klikovi~imaimpresija, data=data1, FUN=siterange)
  imaimpresija Klikovi.FUN1 Klikovi.FUN2 Klikovi.FUN3 Klikovi.FUN4
1  (-Inf,0]          3066           0  0.00000000           0
2   (0, Inf]       455375           0  0.09321768           4
> ggplot(subset(data1, Impressije>0), aes(x=Klikovi/Impresije,colour=katgod))+geom_density()
```



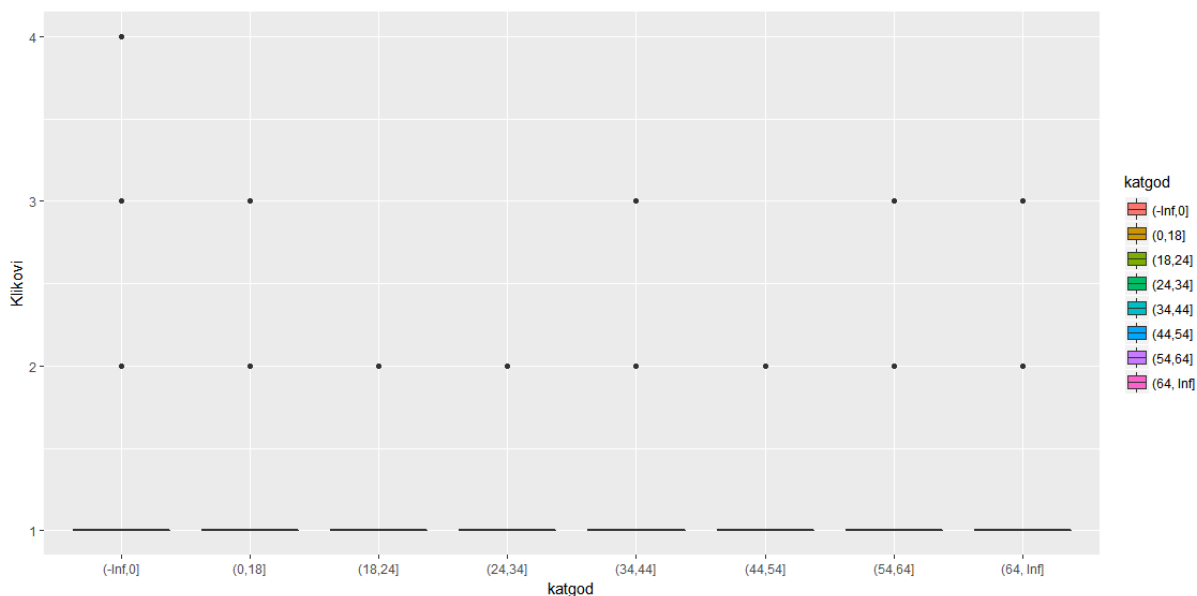
Slika 3.2.3. Density plot / glatka verzija histograma za nyt dataset dan 1 (za podskup podataka gde je broj utisaka>0)

```
> ggplot(subset(data1, Klikovi>0), aes(x=Klikovi/Impresije, colour=katgod))+geom_density()
```



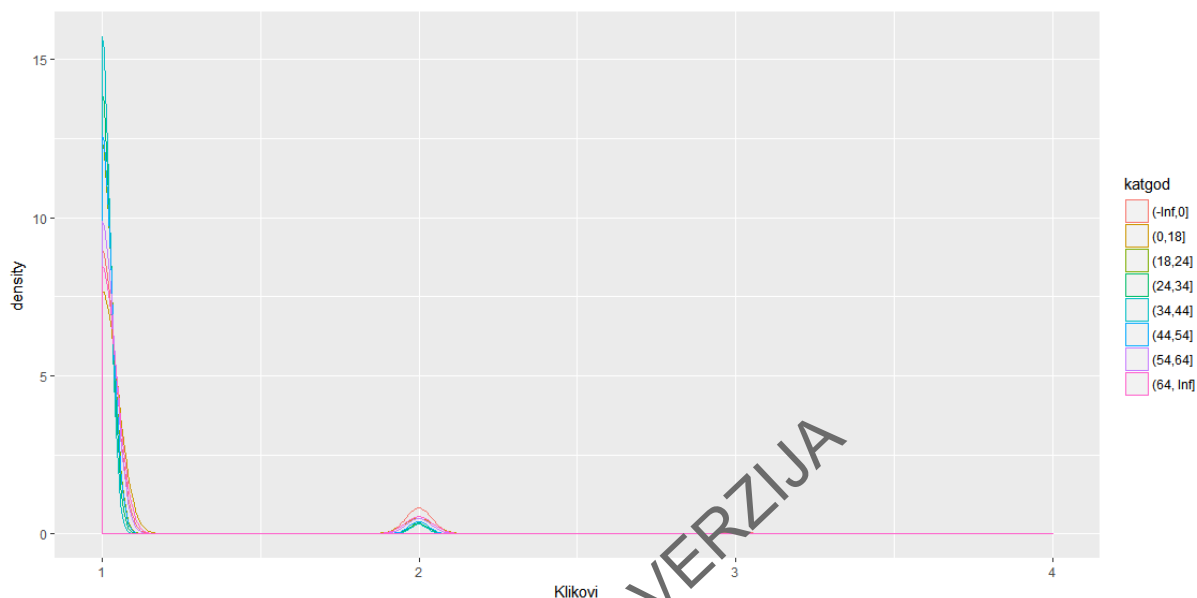
Slika 3.2.4. Density plot / glatka verzija histograma za nyt dataset dan 1 Klikove/Impresije (za podskup podataka gde je broj Klikova>0)

```
> ggplot(subset(data1, Klikovi>0), aes(x=katgod, y=Klikovi, fill=katgod))+geom_boxplot()
```



Slika 3.2.5. boxplot klikova>0 po starosnim kategorijama

```
> ggplot(subset(data1, Klikovi>0), aes(x=Klikovi, colour=katgod)) + geom_density()
```



Slika 3.2.6. Density plot za nyt dataset dan 1 Klikove (za podskup podataka gde je broj Klikova>0)

```
> data1$score[data1$Impresije==0] <- "NemaImp"
> data1$score[data1$Impresije>0] <- "ImaImp"
> data1$score[data1$Klikovi>0] <- "ImaKlikova"
> data1$score <- factor(data1$score)
> clen <- function(x){c(length(x))}
> etabela<-summaryBy(Impresije~score+Pol+katgod,data = data1, FUN=clen)
> etabela
```

	score	Pol	katgod	Impresije.clen
1	ImaImp	0	(-Inf,0]	118401
2	ImaImp	0	(0,18]	6001
3	ImaImp	0	(18,24]	15538
4	ImaImp	0	(24,34]	25690
5	ImaImp	0	(34,44]	31290
6	ImaImp	0	(44,54]	28563
7	ImaImp	0	(54,64]	18626
8	ImaImp	0	(64, Inf]	15585
9	ImaImp	1	(0,18]	10754
10	ImaImp	1	(18,24]	17807
11	ImaImp	1	(24,34]	29241
12	ImaImp	1	(34,44]	35512
13	ImaImp	1	(44,54]	32143
14	ImaImp	1	(54,64]	21499
15	ImaImp	1	(64, Inf]	8887
16	ImaKlikova	0	(-Inf,0]	17776
17	ImaKlikova	0	(0,18]	846
18	ImaKlikova	0	(18,24]	779
19	ImaKlikova	0	(24,34]	1361



```
20 ImaKlikova 0 (34,44] 1675
21 ImaKlikova 0 (44,54] 1494
22 ImaKlikova 0 (54,64] 2006
23 ImaKlikova 0 (64, Inf] 2598
24 ImaKlikova 1 (0,18] 1525
25 ImaKlikova 1 (18,24] 890
26 ImaKlikova 1 (24,34] 1509
27 ImaKlikova 1 (34,44] 1917
28 ImaKlikova 1 (44,54] 1645
29 ImaKlikova 1 (54,64] 2331
30 ImaKlikova 1 (64, Inf] 1486
31 NemaImp 0 (-Inf,0] 929
32 NemaImp 0 (0,18] 43
33 NemaImp 0 (18,24] 124
34 NemaImp 0 (24,34] 165
35 NemaImp 0 (34,44] 219
36 NemaImp 0 (44,54] 224
37 NemaImp 0 (54,64] 118
38 NemaImp 0 (64, Inf] 125
39 NemaImp 1 (0,18] 83
40 NemaImp 1 (18,24] 132
41 NemaImp 1 (24,34] 208
42 NemaImp 1 (34,44] 247
43 NemaImp 1 (44,54] 219
44 NemaImp 1 (54,64] 158
45 NemaImp 1 (64, Inf] 72
> rm(data1)
```

RADNA VERZIJA

### 3.3. Uvoz iz ostalih tipova datoteka

Uvoz iz Excel xlsx fajlova (potrebna java instalacija).

```
> install.packages("xlsx")
> library(xlsx)

> data <- read.xlsx("datafile.xlsx", sheetIndex=1) # podaci sa prvog radnog lista
> data <- read.xlsx("datafile.xls", sheetName="nekiradnilist") #podaci sa imenovanog sheeta
```

Uvoz is SPSS-a :

```
> install.packages("foreign")
> library(foreign)
> data <- read.spss("datafile.sav")
```



**НАПОМЕНА: Ovo je radna (početna) verzija materijala ... pa su greške moguće !!!**



**RADNA VERZIJA**

