**R (uvodno)**

Naredbe u R-u pokreću se kombinacijom tipki Crtl + Enter. Pokreće se naredba iz onog retka u kom je postavljen kursor. Ako želimo pokrenuti više naredbi odjednom, označimo ih mišem i ponovo pritisnemo Crtl+Enter.

Ako ne znamo što radi neka naredba ili koji su joj argumenti, upišemo help(naredba), te se u donjem desnom prozoru prikazuje opis naredbe, često s primjerima korištenja.

Varijable

* definiranje varijabli: npr. a <- 3 ili a=3
* na varijablama možemo vršiti uobičajene operacije

Vektori

* u vektore spremamo elemente istog tipa
* definiranje vektora: npr. x<-c(3, 8, 24)
* **U R-u indeksiranje u vektoru ide od 1, a ne od 0!**
* Dohvaćanje elemenata vektora: npr. x[2] vraća 8, x[c(1,3)] vraća vektor koji na prvom mjestu ima 3, a na drugom 24, x[-c(1,3)] vraća sve elemente vektora x osim na mjestima 1 i 3.
* Posebne vrste vektora:

- i:j – vektor koji se sastoji od cijelih brojeva od i do j (obje granice uključene)

- seq(from=i, to=j, by=k) – vektor koji se sastoji od članova aritmetičkog niza koji počinje s i, te se povećava za korak k sve dok je član niza manji ili jednak j, npr. seq(3, 10, 2) generira vektor c(3, 5, 7, 9)

- seq (from =I, by=k, length.out=n) - vektor koji se sastoji od n članova aritmetičkog niza s korakom k koji počinje od i

- rep() – generira vektor u kom se vrijednosti ponavljaju, npr. rep(“a”, 5) generira vektor c(“a”, “a”, “a”, “a”, “a”), rep(c(1, 3, 5), 3) generira vektor u kom se c(1, 3, 5) ponavlja 3 puta, rep(c(2, 3, 5, 7), c(2, 1, 4, 3)) generira vektor u kom se 2 ponavlja 2 puta, 3 jednom, 5 4 puta, a 7 3 puta

* operacije s vektorima: vektori se mogu zbrajati (+), oduzimati (-), množiti (\*), dijeliti (/), potencirati (^) I provoditi modularnu aritmetiku (%%) element po element
* U R-u su vektori zapravo jednodimenzionalne matrice.
* Naredba x>0 vraća logički vektor (sadrži TRUE ili FALSE) ovisno je li element vektora x na određenom mjestu veći od 0, npr. za x<-c(-1, 7, 3, 0), x>0 vraća vektor c(FALSE, TRUE, TRUE, FALSE), x[x>0] vraća c(7, 3), tj. vektor s onim elementima vektora x koji su veći od 0 u istom poretku kao u originalnom vektoru, which (x>0) vraća pozicije elemenata većih od 0 u vektoru x (u ovom slučaju to bi bile pozicije 2 I 3)
* dodatne funkcije korisne za vektore:

- length(x) – duljina vektora x

- sum(x) – suma elemenata vektora x

- prod(x) – produkt elemenata vektora x

- sort(x, decreasing=FALSE) sortira elemente vektora x uzlazno, a sort(x, decreasing=TRUE) ih sortira u silaznom poretku

- append (x, e, p) – dodaje element e u vektor x iza pozicije p, npr. append(x, 5, length(x)) dodaje 5 na kraj vektora x (iza zadnjeg trenutnog elementa)

Logički operatori

* x==y vraća TRUE ako su x I y jednaki (po vrijednostima koje sadrže), inače vraća FALSE
* x!=y vraća TRUE ako su x I y vrijednosno različiti, inače vraća FALSE
* OR operator (oznaka |), npr. x>0 | y>0 vraća TRUE ako je barem jedan uvjet ispunjen, inače FALSE
* AND operator (oznaka &), npr. x>0 & y>0 vraća TRUE ako su oba uvjeta ispunjena, inače FALSE
* NOT operator (oznaka !), npr. !(x>0) vraća TRUE ako je x>0 FALSE, inače vraća FASLE

Liste

* U liste možemo spremati podatke ražličitog tipa.
* Inicijalizacija liste: npr. lista koja se sastoji od numeričkog vektora, logičkog vektora i stringa
* l<-list(x=c(2, 10), y=c(T, F, T), z=”statistika”). Ovako definirana lista svojim objektima daje nazive: prvi se vektor zove x, drugi y, a treći z.
* Dohvaćanje elemenata liste: l[[i]] vraća i-ti element liste, a ako i-ti element liste ima I naziv ime, možemo ga dohvatiti s l$ime.

Matrice

* Matrice su dvodimenzionalni vektori.
* Svi su elementi matrice istog tipa.
* Inicijalizacija: M<-matrix(data, nrow, ncol, byrow, …), npr. x<- c(1, 2, 3, 4, 5, 6, 7, 8), M<-matrix(x, nrow=4, ncol=2, byrow=T) vraća matricu s 4 retka I 2 stupca popunjenu elementima vektora x popunjavanu slijedno po retcima slijeva nadesno (da smo stavili byrow=F, M bi se popunjavala slijedno po stupcima, odozgo prema dolje)
* Dohvaćanje elementa na presjeku i-tog retka I j-tog stupca: M[i,j], i-ti redak dohvaćamo s M[i, ], a j-ti stupac s M[, j].
* Iz početne matrice možemo generirati bilo koju podmatricu M[r, s], gdje je r vektor indeksa redaka, a s vektor indeksa stupaca, npr. M[1:2, 1:2] vraća prvu 2x2 podmatricu.
* dim(M) vraća dvočlani vektor dimenzija matrice (prva dimenzija se odnosi na broj redaka, a druga na broj stupaca).
* Ponekad želimo na početku alocirati nulmatricu, a zatim po potrebi u nju spremati željene elemente, npr. M<-matrix(rep(0, br\_r\*br\_s), nrow=br\_r, ncol=br\_s) za zadane br\_r I br\_s, te naknadno možemo elementu na poziciji (i,j) pridružiti vrijednost 3.14 s M[i,j]=3.14.
* Na matricama se mogu vršiti iste operacije kao na vektorima. Naredba A\*B množi matrice A I B element po element, dok naredba A%\*%B vraća njihov matrični umnožak.
* Za dodavanje cijelog retka ili stupca u matricu koristimo naredbe rbind() i cbind() redom.

Petlje

* U R-u imamo dvije vrste petlji:
* **for petlja** (v je vektor po čijim elementima prolazi i te se svaki puta jednom izvrši iteracija naredbi napisanih u tijelu petlje, često za v zadajemo vektor 1:n).

for (i in v){

naredbe

}

* **while petlja** (uvjet je neki logički izraz te se zadane naredbe u tijelu petlje izvršavaje dokle god je on istinit):

while (uvjet){

naredbe

}

npr. while (x[i]>0){

i=i+1

}

Data frameovi

* Imaju izgled matrice (tablice) u kojima su tipovi podataka spremljeni po stupcima, dok jedan redak odgovara informaciji o jednoj jedinici.
* Obično se stupcima daju nazivi prema kojima ćemo ih dohvaćati.
* Stupci u data frameovima mogu biti različitog tipa.
* Npr.: želimo li spremiti imena, prezimena, JMBAG-ove studenata koji pohađaju ovaj kolegij, možemo napraviti data frame s tri stupca i redaka koliko je studenata koji su upisani na kolegij.
* Alokacija data framea: npr. studenti<-data.frame (Ime, Prezime, JMBAG), gdje su Ime, Prezime i JMBAG vektori odgovarajućih podataka.
* U R-u je moguće direktno učitavati podatke spremljene na računalu kao data frame (te datoteke mogu imati nastavke .txt, .dat, .csv, .xslx…). Prvi redak u datoteci s podacima može sadržavati informacije o nazivima stupaca – ako je tako, treba u R-u zadati da spremi nazive stupaca.
* Dohvaćanje stupaca: npr. za data frame df u kojem prvi stupac ima naziv Ime, prvi stupac možemo dohvatiti s df$Ime ili df[, 1] (slično kao kod matrica).
* Novi redak (informacije o novoj jedinici) možemo dodati korištenjem naredbe rbind(), a ukoliko želimo postojeći data frame df nadopuniti s još jednim stupcem informacija, možemo napisati naredbu df$naziv\_novog\_stupca<-novi\_stupac.

Funkcije

* Uz standardne matematičke funkcije koje su već definirane u R-u, možemo I sami definirati funkcije.
* Sintaksa za definiciju funkcije:

naziv\_funkcije<-function(ulazni argumenti){

naredbe

return (izlazni argument)

}

* Nakon pisanja funkcije moramo ju pokrenuti (npr. označimo cijelu definiciju funkcije i pristisnemo gumb Run), a zatim je možemo koristiti pomoću naziv\_funkcije(ulazni argumenti) čak i unutar drugih funkcija u R-skripti.
* Defaultni argumenti: ako kao argument u definciji funkcije stavimo x=7, tada se za zadanu vrijednost x-a uzima 7, osim ako korisnik ne zada drugačije prilikom poziva funkcije.

Grafički prikaz podataka

* Naredba plot:

- Ako je y vektor, tada plot(y) crta graf gdje su na x-osi pripadni indeksi, a na y-osi pripadne elementi vektora y.

- Ako su x I y vektori iste duljine, plot(x,y) crta x na x-osi, a y na y-osi.

- Kao glavni argument za naredbu plot također možemo poslati i dvostupčanu ili dvoretčanu matricu.

- Ako želimo na postojeći graf dodati još poneki, imamo sljedeće mogućnosti:

- points(x,y) dodaje točke na postojeći graf

- lines(x,y) dodaje linije na postojeći graf

- abline(a,b) dodaje pravac y=bx+a na postojeći graf

- text(x,y) dodaje tekst na mjesto (x,y)

- curve(f, from=a, to=b) crta graf funkcije f iznad segmenta [a,b].

- Ako želimo prikazati više odvojenih plotova u jednom prozoru, prvo pozovemo naredbu par(mfrow=c(r,s)) koja će podijeliti prozor za plot na r redaka I s stupaca. **VAŽNO:** naredba par() globalno mijenja postavke plota – ona se ne primjenjuje samo na jedan plot. Da bismo u prozoru za plot ponovo imali samo jedan grafički prikaz, možemo zadati par(mfrow=c(1,1)) ili koristiti naredbu dev.off() (koja briše prethodno nacrtane grafičke prikaze iz prozora za plotove) te ponovo nacrtati željeni plot.

- U plotu možemo mijenjati vrstu veze između točaka, nazive osi, naziv plota i sl. (vidi help).

* Za crtanje stupčastog dijagrama, koristimo naredbu barplot(). Glavni argument je vektor frekvencija koje crtamo na y-os. Ako želimo dodati nazive stupića, specificiramo to u argumentu names.arg.
* Za crtanje histograma koristimo naredbu hist(). Glavni argument je uzorak čiji histogtam želimo nacrtati. Naredba hist() ima I neke dodatne argumente, kao npr.

- breaks – vektor granica za razrede histograma, a može biti I (broj+1) razreda u histogramu (ako zadamo breaks=7, histogram će imati 6 razreda, tj. 7 granica)

- probability – TRUE ili FALSE, ovisno o tome želi li se normirani histogram.

Vjerojatnosne funkcije u R-u (detaljnije kasnije)

* Funkcije koje vežemo uz zadanu slučajnu varijablu X su njena funkcija distribucije, funkcija gustoće, generalizirani inverz funkcije distribucije I slično.
* Vjerojatnosne funkcije u R-u označene su sljedećim slovima (koja se dodaju ispred skraćenog imena distribucije slučajne varijable X):

- p – simbol za funkciju distribucije FX. Glavni argument je točka x u kojoj računamo vrijednost FX(x).

- d – simbol za funkciju gustoće fX. Glavni argument je točka x u kojoj računamo vrijednost fX(x).

- q – simbol za generalizirani inverz funkcije distribucije. . Glavni argument je realni broj 0<a<1 te funkcija vraća najmanji realni broj x takav da je a<=FX(x), tj. vraća a-ti kvantil.

- r – simbol za generiranje uzorka. Glavni argument je veličina uzorka n.