

FUNKCIJE U R-U

STATISTIČKI PRAKTIKUM 2

15. VJEŽBE

Neki dodatni paketi u R-u

1. dplyr
2. tidyr
3. purr
4. lubridate
5. stringr
6. recipes
7. ggplot2

DPLYR

- ▶ noviji paket (2014.)
- ▶ sadrži funkcije za manipulaciju podacima
- ▶ podešen za rad s okvirima podataka
- ▶ podaci mogu biti spremljeni u vanjskim bazama
- ▶ najvažnije funkcije:
 - `select()` .. odabir stupaca
 - `filter()` .. odabir redaka
 - `mutate()` .. dodavanje novih stupaca
 - `group_by()` .. grupiranje
 - `summarize()` .. analiziranje podataka po grupama

GGPLOT2

- ▶ funkcije za grafički prikaz podataka
- ▶ također radi s okvirima podataka
- ▶ funkcija `qplot()` stvara prazan graf na koji se dodaju elementi (slojeve spajamo oznakom `+`)
- ▶ elementi grafa:
 - ▶ geometrijski objekti: `geom_point()` (točke), `geom_line()` (linije), `geom_boxplot()`, `geom_bar()`, `geom_histogram()`
 - argumenti koji se odnose na povezivanje podataka se unose unutar podnaredbe `aes()` (npr. `color ..` boja ovisno o nekoj drugoj varijabli), ostale izvan (npr. `size ..` veličina svih točaka)
 - ▶ dijelovi grafa za uređivanje, nisu novi podaci: `theme()`
 - `axis.title`, `axis.text.x`
 - ▶ ostali elementi koji se dodaju na graf
 - `ggtitle()` (naziv grafa), `xlab()` (naziv x-osi)

Zadatak 1

Učitajte podatke naziva `mtcars` o svojstvima i performansama 32 marke automobila.

- (a) Izdvojite samo podatke o potrošnji (`mpg`), broju konjskih snaga (`hp`) i vrsti prijenosa (`am`).
- (b) Ispišite one marke automobila čiji broj cilindara (`cyl`) je jednak 8, a potrošnja nije veća od 16.
- (c) Spremite u novu varijablu podatke o potrošnji, broju cilindara, težini (`wt`), vremenu za 1/4 milje (`qsec`) i vrsti prijenosa za one automobile koji čiji oblik motora je ravan (`vs` je jednak 1).
- (d) U tu novu matricu dodajte novi stupac u kojem ćete spremiti podatke od potrošnji po jednom cilindru (`mpg/cyl`).
- (e) Izračunajte prosječnu potrošnju automobila ovisno o broju cilindara, a zatim minimalno i maksimalno vrijeme za 1/4 milje (`qsec`) ovisno o broju cilindara i vrsti motora (`cyl` i `vs`).

Zadatak 2

Za iste podatke napravite sljedeće:

- (a) Grafički prikažite odnos između potrošnje i vremena za 1/4 milje. Neka oblik točke sugerira broj cilindara pojedinog automobila, a boja opisuje težinu. Povećajte veličinu točaka na 4.
- (b) Dodajte naziv grafa i osi. Povećajte naslov grafa.
- (c) Nacrtajte histogram podataka o potrošnji. Promijenite boju grafa i zarotirajte oznake na x -osi za 60 stupnjeva.
- (d) Nacrtajte boxplot podataka o potrošnji ovisno o vrsti motora. Svaku grupu obojite drugom bojom.