

Zadatak: Težina

Izvor: <http://lib.stat.cmu.edu/datasets/bodyfat>

- a) Na temelju varijabli *Weight* i *Height* izračunajte novu varijablu *BMI* (Body mass index).
- b) Opišite varijable statističkog skupa (uključujući i varijablu definiranu u a)) i grafički ih prikazite.
- c) Izračunajte matricu korelacija između svih varijabli i interpretirajte rezultate. Grafički prikazite matricu korelacija.
- d) Ispitajte normalnost razdiobe za promatrane varijable.
- e) Formirajte novu varijablu *BMI_faktor* rekodiranjem varijable *BMI* na sljedeći način:

<20	Pothranjenost
20-25	Idealna težina
25-30	Prekomjerna tjelesna masa
>30	Pretilost.
- f) Ispitajte postoji li razlika u varijabli *Percent body fat*, *Neck circumference*, *Chest circumference*, *Abdomen 2 circumference*, *Hip circumference*, *Thigh circumference*, *Knee circumference*, *Ankle circumference*, *Biceps (extended) circumference*, *Forearm circumference* i *Wrist circumference* u ovisnosti o modalitetima varijable *BMI_faktor* korištenjem jednofaktorske analize varijance. Ispitajte pretpostavke za primjenu analize varijance. Ako postoje razlike promatranih varijabli po modalitetima varijable *BMI_faktor* ispitajte između kojih modaliteta korištenjem post hoc testa. Ako nisu ispunjene pretpostavke za provedbu analize varijance provedite testiranje i korištenjem odgovarajućeg neparametarskog testa.
- g) Definirajte model regresije kod kojeg će zavisna varijabla biti *Percent body fat*, a nezavisne varijable ostale varijable u originalnom skupu podataka (bez varijable *BMI*) te varijabla *BMI_faktor*. Komentirajte parametre regresije: koeficijent determinacije i korigirani koeficijent determinacije. Interpretirajte skupni i pojedinačne testove signifikantnosti regresije za svaku od promatranih nezavisnih varijabli. Komentirajte koeficijente u jednadžbi regresije. Provedite izbor varijabli koristeći neku od metoda za izbor varijabli. Nacrtajte normalni prikaz rezidualnih vrijednosti za regresijski model i interpretirajte ga.