

DZ 2

STATISTIČKI PRAKTIKUM 2

20. listopada 2023.

Zadatak 3. - Cijena nekretnina

U datoteci `nekretnine.txt` nalaze se podaci vezani uz cijenu nekretnina. Stupci su redom: veličina nekretnine u m^2 (Kvadr), broj spavaćih soba (Br_soba), godina izgradnje nekretnine (God_gradnje), udaljenost od centra grada u kilometrima (Udalj_centra), cijena nekretnine izražena u tisućama dolara (Cijena), tip nekretnine, pri čemu "S" označava stan, a "K" kuću (Tip) te broj prethodnih vlasnika nekretnine (Br_vlasnika).

- (a) Grafički i inferencijalno testirajte statističku ovisnost cijene o tipu nekretnine. Postoji li značajna razlika u cijeni nekretnine ovisno o njenom tipu?
- (b) Odredite prosječnu kvadraturu za stanove te za kuće. Statističkim testom ispitajte ima li razlike između veličine nekretnine obzirom na tip.

- (c) Ispitajte linearnu povezanost varijabli $Kvadr$ i $Udalj_centra$. Možemo li tvrditi da kvadratura nekretnine ne ovisi o njenoj udaljenosti od centra grada?
- (d) Grafički, a zatim višeparametarskim linearnim modelom, istražite ovisnost cijene nekretnine o kvadraturi, broju soba, godini gradnje i njenoj udaljenosti od centra.
- (e) Interpretirajte taj model te komentirajte kriterije za kvalitetu modela. Je li moguće poboljšati model? Kako biste, ako je moguće, to napravili?
- (f) Ako nas zanima stan koji je izgrađen je prije 3 godine, nalazi se 2.1 km od centra te ima 93 kvadrata, s dvije spavaće sobe, procijenite očekivanu cijenu zajedno s pripadnim 90% pouzdanim intervalom.

- (g) Na razni značajnosti od 10% testirajte hipotezu da koeficijent uz godinu gradnje treba biti jednak 3, a koeficijent uz udaljenost od centra jednak 1.
- (h) Od sada nadalje promatramo model koji nam opisuje linearnu ovisnost cijene nekretnine o njezinoj veličini (kvadraturi). Grafički i inferencijalno provjerite pretpostavke tog linearnog modela (proučite možete li iskoristiti Durbin-Watsonov test iz paketa (`lmtest`) te Jarque-Berov test). Ima li u tom modelu utjecajnih točaka? Ako ima, kako bi na naš model utjecalo uklanjanje tih točaka?
- (i) Nekim od kriterija usporedbe modela provjerite je li ovaj model s jednim parametrom bolji od početnog modela kojeg smo promatrali.

- (j) Grafički prikažite podatke o cijeni nekretnina, zajedno s 80% pouzdanom prugom za opažanja.
 - (k) Također za dobiveni model odredite 90% pouzdane intervale za procijenjene parametre.
 - (l) Ispitajte ovisnost cijene nekretnine o kvadraturi i odredite ovisi li taj omjer o tipu nekretnine. Interpretirajte tu ovisnost uz odgovarajuće grafičke prikaze i statističke testove.
- (40 bodova)