

Navodila za seminarsko nalogo iz linearne regresije

1 Uvod

Narediti morate statistično analizo podatkov in interpretirati dobljene rezultate. Rezultat vaše naloge naj bo poročilo v formatu PDF, narejeno na osnovi Rmd vzorca **linregresija_vzorec.Rmd**, ki naj vsebuje kodo, spremno besedilo in rezultate ter grafe. Kodo naj spremlja besedilo, ki razloži rezultate, opiše, kako do rezultatov pridemo in pove, kaj sploh računamo.

2 Podatki

Katere podatke (.csv) morate uporabiti, določite na osnovi začetne črke vašega priimka:

Priimek	Podatki	Opis
A – F	jezero.csv	Dolžina ribe
G – K	mozgani.csv	Teža možganov pri sesalcih
L – P	klek.csv	Višina kleka
R – Š	zavor.csv	Zavorna pot
T – Ž	forbes.csv	Zračni tlak

Podatki so bolj podrobno opisani v dokumentu **VISpodatki.pdf**.

3 Podrobnejša navodila

V dokumentu **Linregresija.pdf** boste našli bolj podroben opis metode linearne regresije. Seminarska naloga mora vsebovati naslednje dele (v opisanem vrstnem redu).

1. Opis podatkov
2. Opisna statistika originalnih podatkov
3. Razsevni diagram in vzorčni koeficient korelacije
4. Formiranje linearnega regresijskega modela, enačba vzorčne regresijske premice
5. Točke visokega vzvoda in osamelci
6. Preverjanje predpostavk linearnega modela (diagnostični grafi in njihova obrazložitev)
 - (a) Linearnost modela
 - (b) Normalnost porazdelitve naključnih napak
 - (c) Homogenost variance: graf in Breusch-Paganov test
 - (d) Cookova razdalja: graf in analiza vpliva točk preko pogoja velikega vpliva točk.
7. Testiranje linearnosti regresijskega modela in koeficient determinacije
8. Intervala zaupanja za naklon in odsek regresijske premice
9. Interval predikcije za bodočo vrednost slučajne spremenljivke Y pri izbrani vrednosti X (za vsaj tri vrednosti)
10. Zaključek.

Analiza v točkah 3-9 se dela za transformirane podatke.

Bonus točka Po opisni statistiki podajte razloge za transformacijo podatkov na osnovi razsevnega diagrama, koeficienta korelacije in diagnostičnih grafov originalnih (netransformiranih) podatkov. Potem pa nadaljujete z analizo v točkah 3-9.