

Metodologija znanstveno-istraživačkog rada s prostornim podacima u okruženju za statističko računarstvo R

Plan seminara

Voditelj seminara: prof. dr. sc. Renato Filjar

Prethodno predznanje: osnove linearne algebre.

Potrebna oprema: polaznici su savjetovani uzeti vlastito osobno računalo, neovisno o vrsti OS-a, radi instalacije besplatne open-source programske podrške, potreban pristup internetu za pristup instalacijskim datotekama i repozitorijima podataka u otvorenom pristupu

Programska podrška koja će biti instalirana: R ver. 4.0.2 (ili noviji), Rstudio v. 1.3.1073 (ili noviji), dodatne R knjižnice

Sadržaj seminara (svaka tema pokriva jedan sat, 60 min, problemski orijentirane kombinirane teorijsko-praktične nastave)

DAN 1

1. Uvod i motivacija. Instalacija programskog okruženja R i grafičkog korisničkog sučelja RStudio. Organizacija znanstveno-istraživačkog rada. Prepoznavanje problema – pregled prethodnih istraživanja – oblikovanje hipoteze – osmišljavanje metodologije rada – izbor i prikupljanje eksperimentalnih podataka – opisna statistička analiza podataka – razvoj modela – vrednovanje modela – primjena modela. Vrste modela.
2. Podatkovne strukture u R-u.
3. Priprema dokumentacije u R-u (DOCX, PDF, HTML, korištenjem *RMarkdown* i *LaTeX* programskih alata).

DAN 2

4. Grafički prikaz u okruženju R I. Osnove korištenja instrukcije *plot()*.
5. Upravljanje podacima u R-u. Rad s podatkovnim okvirima (*data frames*).
6. Opisna statistička analiza podataka u R-u. Aritmetička sredina, kvartile, standardna devijacija, funkcija gustoće vjerojatnosti, funkcija kumulativne vjerojatnosti, kutijasti dijagrami, Q-Q dijagrami.

DAN 3

7. Statistički testovi u R-u. Shappiro-Wilk, Kolmogorov-Smirnov, t-test, F-test.
8. Grafički prikaz u okruženju R II. Procjena eksperimentalne statističke razdiobe.
9. Rad s vremenskim nizovima u R-u. Osnovna analiza vremenskog niza. Dekompozicija vremenskog niza. Prognoza vremenskog niza.

DAN 4

10. Razvoj modela zasnovanog na podacima u okruženju R. Vrednovanje modela analizom reziduala (R^2 , $adj R^2$, $RMSE$, $P-O$ dijagrami). Jednostruka linearna regresija.
11. Višestruka linearna regresija. Mješovita linearna regresija.
12. Prostorni podatci i njihovi formati. Rasterski i vektorski prostorni podatci. Pristup otvorenim bazama prostornih podataka pomoću okruženja R.

DAN 5

13. Referentni koordinatni sustavi. Karte. Grafički prikaz prostornih podataka u okruženju R. Prikaz statističkih prostornih podataka u okruženju R. Prikaz putanja u okruženju R.
14. Osnove prostorne statističke analize u R-u. Skale i udaljenosti. Prostorna korelacija. Prostorna interpolacija. Lokalna regresija.
15. Osnovne prostornog statističkog modeliranja u R-u. Interpolacija metodom težinskih koeficijenata. Kriging.
16. Provjera ovladavanja predstavljenim materijalom. Izdavanje potvrda o sudjelovanju.

Literatura:

1. Hediger, M. (2020). The R introductory course. University of Zurich. Zurich, Switzerland. Dostupno na: <https://www.math.uzh.ch/typo3conf/ext/qfq/Classes/Api/download.php?s=5f2e608d191fc>
2. Lilja, D J. (2016). Linear Regression Using R: An introduction to data modelling. University of Minnesota. Dostupno na: <https://conservancy.umn.edu/handle/11299/189222>
3. Dettling, M. (2019). Applied Time Series Analysis. ETH Zurich. Zurich, Switzerland. Dostupno na: <https://polybox.ethz.ch/index.php/s/F1yrwqQ8YQwiXON/download>
4. Filić, M. (2017). Analiza postupka procjene položaja temeljem zadanih pseudoudaljenosti u programski određenom prijamniku za satelitsku navigaciju (diplomski rad). Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet, Matematički odsjek. Zagreb, Hrvatska. Dostupno na: <https://zir.nsk.hr/islandora/object/pmf%3A3230>
5. Hijmans, R. (2020). Spatial Data Science with R. University of California at Davis. Davis, CA. Dostupno na: <https://rspatial.org>