

Pokročilá letní škola: Techniky k vizualizace dat (termín předběžně 5.-8.9.2022)

Den 1. Vizualizace v R (Lukáš Hájek)

Základní grafické nástroje v oblasti jednorozměrné statistiky. Přehled grafů a jejich využití. Složitější grafické nástroje pro analýzu dat. Úvod práce s R a R Studio. Základní grafické nástroje v R. Ggplot jako oblíbený balíček pro vizualizaci. Chyby ve vizualizacích dat.

Den 2. Mnohorozměrné škálování (Petr Soukup)

Jedno a mnohorozměrné škálování – základní logika. Úlohy řešitelné skrze mnohorozměrné škálování (MDS). Měření vzdáleností v mnohorozměrném prostoru pro různé typy proměnných. Matice vzdáleností a podobností. Logika MDS jako redukce mnohorozměrného prostoru. Základní algoritmy MDS. Demonstrace MDS na fyzické mapě. Využití MDS pro reálné analýzy sociálněvědních dat. MDS jako alternativa k EFA.

Den 3. Korespondenční analýza. (Petr Soukup)

Možnosti užití klasické analýzy kontingenčních tabulek pro více než 2 proměnné. Načítání kontingenční tabulky do datové matice a práce s ní. Korespondenční tabulka. Řádkové a sloupcové profily, inercie a míry jejího vysvětlení. Základní korespondenční analýza pro dvě kategoriální proměnné. Vyhodnocení korespondenční analýzy – grafické a tabulkové. Interpretace výsledků. Korespondenční analýza pro více než dvě kategoriální proměnné. Vyhodnocení a interpretace výsledků.

Den 4. Klasifikační a regresní stromy. (Petr Soukup)

Základní logika regresního přístupu, motivační příklad. Lineární a nelineární model, model logistické regrese. Klasifikační a regresní stromy (CART) jako alternativa ke klasickému regresnímu přístupu. Rozdíl mezi klasifikačními a regresními stromy. Jednotlivé algoritmy – CHAID, exhaustive CHAID, CRT, QUEST aj. Využití CART na sociálněvědních datech.

Předpoklady účasti: Letní škola je určena pro studenty (Mgr, a Ph.D.) i pro akademické pracovníky či výzkumníky. Účastník by měl být seznámen s postupy popisné statistiky, s korelační analýzou a se základními postupy při využití statistického testování a intervalů spolehlivosti. Výhodou je schopnost pracovat v software (pro demonstrace bude užito R, JAMOVI a SPSS).

Účastník by měl být seznámen s postupy popisné statistiky, s korelační analýzou a se základními postupy při využití statistického testování a intervalů spolehlivosti. Výhodou je schopnost pracovat v software.

Průběh letní školy: každý den bude výuka probíhat v čase 9-12 a 13-16. Z hlediska formy půjde o kombinaci přednášky a praktického cvičení. Účastníci budou mít k dispozici prezentaci (pdf) a dále datové soubory a případně skripty pro analýzy v software. Na konci každého dne bude blok vyhrazený k samostatnému procvičení konkrétní techniky.

Cena pro účastníka: 4 tisíce Kč (v ceně je kurzovné, materiály a drobné občerstvení). Možnost účasti jen ve vybraný den (vybrané dny) po dohodě s organizátory, v tomto případě bude cena stanovena individuálně.

Místo konání: FF UPOL, Olomouc, účast je možná i online

Přihlášky: <https://forms.gle/zp5aAikxbsNYm8Xk9>