

Oponentní posudek bakalářské práce

Jméno a příjmení autora práce: Eva Pirohová

Název práce: Popisná statistika v R s aplikací na reálná data

Autorka si pro svoji práci zvolila téma popisné statistiky v R, které je v současné době aktuální. V úvodu práce je stanovený jediný cíl, a to přiblížit práci s R v oblasti popisné statistiky. Tento cíl je velmi obecný, což znesnadňuje ohodnocení, do jaké míry byl splněn. Vhodnější by bylo definování více dílčích cílů. Celkově se práce profiluje jako jakýsi návod, což není příliš vhodné pro bakalářskou práci, která by měla být spíše prací vědeckou.

Práce je rozdělena do tří kapitol. V první kapitole jsou představeny nejnужnější základy práce s R, ve druhé vybrané metody popisné statistiky v R a ve třetí aplikace na reálných datech. Takové členění považuji za logické, vyvstává ale otázka, jestli je pro účely práce první kapitola zapotřebí. V jiných zdrojích je úvod práce s R mnohem lépe a detailněji vysvětlen. Druhá kapitola práce je nejpřínosnější. Autorka v ní představuje funkce, kterými lze vypočítat základní popisné statistiky. Funkce některých statistik sama naprogramovala, ať už z důvodu nepřítomnosti těchto funkcí v R nebo z důvodu zobecnění základních funkcí.

Největší výtky mám ke třetí kapitole, která měla být praktickou aplikací metod představených v kapitole 2. Bohužel, výsledek mi přijde jako mírně rozšířená druhá kapitola, včetně duplikování některých syntaxí. Jsou zde sice představena reálná data, ale bez jakéhokoli výzkumného cíle. Jediné, o co jde, je aplikovat danou metodu představenou v kapitole 2 na datovém souboru, téměř bez interpretace výsledků. Také bych doporučil skript použitý ve třetí kapitole dát až na její konec nebo do přílohy a v textu se věnovat pouze analýze.

Z formálního hlediska má práce dobrou úroveň, literatura je správně citovaná, vzorce číslovány, texty syntaxí a názvy funkcí a balíčků jsou v textu dobře odlišeny. Z jazykového hlediska mi práce přijde mírně nedoladěná. Vyskytuje se v ní poměrně velké množství překlepů a některé formulace působí kostrbatě. Také je zde časté střídání první osoby jednotného a množného čísla a trpného rodu.

S přihlédnutím ke všem výše uvedeným faktům, bakalářskou práci doporučuji k obhajobě a hodnotím ji jako *dobrou*.

Dotazy k obhajobě:

- 1) U koeficientů šikmosti a špičatosti je poměrně časté, že každý software používá mírně odlišný vzorec. Můžete ověřit, zda funkce v balíčku *moments* odpovídají vzorcům 2.14 a 2.15?
- 2) Na str. 39 píšete, že je FD pravidlo robustnější vůči odlehlým hodnotám oproti ostatním pravidlům, přitom jich histogram 3.5. založený na tomto pravidlu obsahuje mnohem více. Můžete to vysvětlit?

V Praze dne 29.5.2017

Ing. Zdeněk Šulc, Ph.D.