



POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Jméno studenta: Vilém Procházka

Název práce: Nástroje vizualizace dat pro podporu rozhodování

Autor posudku: Jiří Haviger

Cíl práce: *citace: Cílem této práce je vymezení pojmu vizualizace dat pro účel této práce, která se bude zabývat možnostmi, jak lze metodami vizualizace dat přispět k porozumění dějů a souvislostí, které jsou v datech uloženy. Východiskem práce bude studium a popis základních principů a nezbytných znalostí, potřebných pro práci s daty za účelem jejich vizualizace, čeho se při vizualizaci vyvarovat a uvést best practices k jednotlivým typům grafů. Dále bude proveden výběr a porovnání programovacích jazyků, které se využívají pro vizualizaci dat. Podrobněji budou porovnané jazyky R a Python. V jednom vybraném jazyce poté bude vytvořena aplikace, která se bude zabývat statickými grafy a úpravami pro přidání interaktivních prvků.*

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vymezení cíle a jeho naplnění	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Zpracování teoretických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zpracování praktických aspektů tématu	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Formální úprava a náležitosti práce	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Vlastní přínos studenta	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Dílčí připomínky:

Téma práce je vhodné a v době masového rozšíření dat potřebné. Zpracování však vykazuje značné nedostatky. Uvedu jen některé:

Název práce neodpovídá obsahu - v názvu je uvedena podpora rozhodování, v práci zazněla pouze v jedné podkapitole jedné konkrétní aplikace.

Cíl práce (kapitola 1.2) není dobře formulován - **není zřejmé, co je hlavním cílem.** Jestli vizualizace dat z průběhu nákazy, nebo metoda vizualizace dat kde data z nákazy

jsou pouze ukázková, nebo výukový materiál pro studium vizualizací dat, kde závěrečné vizualizace jsou opět pouze ukázkové. Konkrétněji k cíli práce:

- Formulace jsou obecné a vágní, např. „přispět k porozumění“
- není jasná operacionalizace tudíž je problematické zhodnotit, zda bylo cílů dosaženo, např. na základě čeho „bude proveden výběr a porovnání programovacích jazyků“ – teoretických poznatků, nebo vlastní porovnání?
- vymezení pojmu jen „pro účely práce“ je nevhodné, práce by měla mít potenciál přenosu poznatků mimo samotnou práci,
- „Východiskem práce bude ... čeho se při vizualizaci vyvarovat a uvést best practices k jednotlivým typům grafů. Tohle je patří do východisek pro práci nebo se ujedná již o cíl práce?
- práce se zabývá nástroji vizualizace, ale ne jejich schopnosti „přispět k porozumění dějů a souvislostí“. Porozumění dějů se těžko ověřuje, proto mi nepřijde vhodné je uvádět v cíli práce.

Protože **hlavní cíl není jasný**, je problematické další ohodnocení práce.

Literární rešerše (kapitola 2) působí jako výpisky z různých zdrojů, u kterých není jasný ani způsob výběru (náhodně?), ani vztah k cíli práce (viz špatně definovaný cíl). **U každého z příspěvků by tedy mělo být uvedeno, jak jeho závěry přispívají k cíli.** Např. jak „2.10 Interaktivní simulační a vizualizační nástroj pro analýzu povodňových událostí použitelný v praxi“ pomůže k vizualizaci dat z nákazy, nebo jak je v analýze využít „2.14 Payola: kolaborativní framework pro analýzu a vizualizaci propojených dat“. Vzhledem k tomu, že práce pojednává o vizualizacích mi v této části **chybí grafy** z uvedených prací.

Analýza (kapitola 3) v části nazvané analýza bych očekával nějakou analýzu, nikoli popis, jak se má analýza dat provádět. Jinak je zde stručně popsán postup analýzy dat, což je vzhledem k rozsahu práce pochopitelné. Výhrady však mám k popisu nástrojů pro vizualizaci dat (3.1.3) – zde jsou **zcela vynechány existující vizualizační nástroje využívané v rámci podpory pro rozhodování a obecně BI** – MS Excel výrazně se rozvíjející podpory vizualizací, Tableau, Qlick a další, viz např Gartner – analytics and business intelligence platforms. Následuje stručný výčet grafů, jejich vlastností a obecná pravidla pro vytváření grafů. **I zde chybí vztah jednotlivých subkapitol k cíli práce.**

Návrh řešení a implementace (kapitola 4). Výběr mezi R - Python autor zdůvodnil. V úvodní části str2-3 ale popisuje výhody jazyka Julia i možnost implementace knihoven z R. **Proč tedy nebyl využit jazyk Julia?** Uvedení 5x jednoho vylepšovaného grafu (každý na 1/2 strany) považuji za mrhání místem. Na grafu str 43 není zřejmá využitá barevná škála, ale i kdyby tak je shodná s osou y – čím vyšší y, tím světlejší odstín. Autor tedy mapuje jednu informaci na dva parametry grafu, což je nadbytečné. Graf 7 je zřejmě chybně popisek „7 Počet nakažených vůči počtu testů za den“ – ukazuje na poměrové číslo, hodnoty by tedy měly být v rozmezí 0-1. Částečně je to vysvětleno v následujících grafech, ale i zde bych doporučil relativní škálu – z grafu není na první pohled zřejmé, že počet nakažených ku počtu testovaných za jeden den klesá. Graf 13 – opět mapování jedné informace na dva parametry grafu – nadbytečné. Graf 16 – využívání koláčových grafů se nedoporučuje, ačkoli s popisky to lze akceptovat. Uspořádání v koláčovém grafu také není vhodné, využívá se od největšího k nejmenšímu (zde Rakousko – Itálie prohozené).

Celkové posouzení práce a zdůvodnění výsledné známky:

Cíl práce není zřejmý a z toho plyne větší množství nedostatků – není zřejmé, proč jsou některé pasáže obsažené a jak se vztahují k cíli práce. Z větší části má práce charakter kompilátu z několik (ocitovaných) zdrojů. Přínos práce je z větší míry pro autora (seznámení s novým tématem). Informatických dovedností nebylo příliš využito. Obsahově práce přesahuje standardní obsah učiva, student prokázal schopnost samostatného studia problematiky vizualizací dat.

Kontrola antiplagiátrským systémem ukazuje shodu 0%

Otázky k obhajobě:

1. Porovnání jazyků R a Python provádíte pouze teoreticky – ukažte a porovnejte kódy ke grafu 5 v Python a v R, případně v Julia
2. Porovnejte výhody a nevýhody existujících vizualizačních nástrojů (Excel, Tableau) oproti programovacím jazykům R/Python/GoogleChart.

Práci doporučuji k obhajobě.**Navržená výsledná známka: E****V Hradci Králové, dne 8. května 2020**

podpis