



POSUDEK VEDOUcíHO DIPLOMOVÉ PRÁCE

Jméno studenta: Natálie Rakovská

Název práce: Porovnání kombinovaných nástrojů pro vizualizaci a analýzu dat

Autor posudku: Hana Skalská

Cíl práce: Kombinované nástroje vizualizace a analýzy dat. Typy dat a možnosti jejich vizualizace a explorační. Volba reprezentantů pro porovnání a zdůvodnění volby, stanovení kritérií pro porovnání SW. Praktické porovnání zvolených reprezentantů.

Povinná kritéria hodnocení práce	Stupeň hodnocení (známka)					
	A	B	C	D	E	F
Práce svým zaměřením odpovídá studovanému oboru	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input 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 checked="" type="checkbox"/>	<input type="checkbox"/>	<input 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 checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Adekvátnost použitých metod, způsob jejich použití	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hloubka a správnost provedené analýzy	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Práce s literaturou	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Logická stavba a členění práce	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jazyková a terminologická úroveň	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input 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 checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Využitelnost výsledků práce v teorii (v praxi)	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Vyjádření k výsledku anti-plagiátorské kontroly

V repozitáři Odevzdej.cz nenalezena shoda (0%).

Dílní připomínky a náměty:

V práci, která není psaná v AJ, je z neznámého důvodu používán anglický termín self-service pro kombinované nástroje analýzy. Správnější termín self-service BI je vysvětlený až v kapitole 6 (vhodný by byl také termín v jazyku práce „kombinované nástroje“).

V úvodní části chybí vysvětlení self-service (výskyt už na prvním řádku úvodu), následně chybí i vysvětlení BI, v dalším také vysvětlení self-service BI. V přehledu zkratk nejsou uvedené všechny zkratky (např. CRM).

Úvodní kapitoly jsou obtížně srozumitelné. Častým prohřeškem jsou duplikovaná souvětí obdobného významu, jen jinak formulovaná (str.1, str. 7 až 17), text není dobře logicky provázaný, některé formulace jsou nepromyšlené až zavádějící.

Str.7 až 9 např.: Management zákazníků – „predikce vlivu reklamy na produkci“, „způsob sběru dat záleží na jejich formátu“, „strukturovaná data způsobila revoluci v systémech

založených na papírových dokumentech...“, „pohled na zákazníka z ptačí perspektivy“, „nestrukturovaná data jsou klíčem k prediktivním analytickým softwarům“, atd.

Definice Big data (str. 9) není aktuální, práce uvádí 3V, chybí + Veracity + Value.

Tabulka 1 prezentuje jiné typy dat, než které jsou v textu příslušné kapitoly popsány. Vysvětlení návaznosti chybí.

Typy dat a měrné stupnice patří mezi základní pojmy při vizualizaci a analýze dat. Autorka DP nepoužívá obvyklé rozlišení a směšuje různé pohledy na data. Vychází z internetových popisů, nedaří se logická návaznost textů. Neuvažují se transformace a předzpracování dat, což by v textu usnadnilo návaznost na další kroky a ujasnilo by popis různých typů a formátů dat pro analýzu a vizualizaci.

Chybí logická provázanost kapitoly 4 (data) s kapitolou 5 (vizualizace), která v částech 5.1 a 5.2 obsahuje spíše nesourodé poznámky, než věcný text. Obrázek 2 (Datová vizualizace) není srozumitelný (v textu není vysvětlený).

Na obrázku 7 není uvedený rok, pro který byl „Magický kvadrant“ sestavený. Lídrem je Microsoft, ne samotný MsPower BI a MsOffice, jak je uvedeno v textu (ale také Azure).

Celkové posouzení práce a zdůvodnění výsledné známky:

Úvodní kapitoly vycházejí převážně z webových zdrojů sice řádně citovaných, ale některé uváděné poznatky a definice jsou převzaté bez posouzení jejich odborné kvality, aktuálnosti, nebo vzájemné konzistence (zejména kapitoly „Statistické typy dat, dělení dat, typy dat“ a „Význam vizualizace dat“). Text obsahuje neodborná vyjádření a nejasná vysvětlení. Stěžejní termín self-service BI je částečně vysvětlený až v kapitole 6, v práci je často uváděný matoucí termín self-service.

Teoretická část práce není myšlenkově sjednocená. Jednotlivé prvky jsou rozptýlené v několika úvodních rešeršních kapitolách a nejsou zřetelně shrnuté. Studentka neposkytla časovou rezervu k pročetění práce před finalizací a včasným připomínkám.

Lépe je propracovaná kapitola 6 (BI a self-service BI) a praktická část (kapitola 7), ve které je uplatněná metodologie CRISP pro popis a vizualizaci dat. Aplikace pro rozsáhlá data porovnává možnosti Tableau a MsPower BI.

Cenný by byl návrh vlastních (uživatelských) kritérií a jejich užití pro porovnání obou zvolených typů softwaru.

Otázky k obhajobě:

Jaká rizika mohou být spojená s odstraněním záznamů s chybějícími hodnotami?

Jaký významový rozdíl v IT komunitě je mezi termíny self-service a self-service BI?

Práci doporučuji k obhajobě.

Navržená výsledná známka: C

V hodnocení mezi C až D se vzhledem k praktické části přikláním k C.

V Hradci Králové, dne 28. prosince 2020

podpis