

Cirkulárna vizualizácia dát dochádzky za prácou na príklade obcí Bratislavského samosprávneho kraja

Vladimír BAČÍK, Dominika CHVOSTEKOVÁ

Circular visualization of commuting data on the example of the Bratislava self-governing region communes

Abstract: Circular visualization has recently become a kind of imaginary leader in the field of graphic visualization, mainly due to its attractive circular layout and the ability to display large amounts of data. In our paper, we focused only on a small subset of this group, namely the creation of chord diagrams that are useful in depicting the interrelationships between the observed regions. Commuting data represent an ideal example of such use, as we can clearly identify the source and target regions. In our paper, we used the communes of the Bratislava self-governing region as a model territory and data on daily commuting to work and schools from 2011. We used two basic tools to create these diagrams. Using Circos online application we created static chord diagrams, interactive diagrams were created using D3 library, which has integrated functions for creating such graphical outputs. It is a relatively frequently used graphic expression in the world, but very rarely in Slovakia and neighboring countries. In this paper we want to point only to the basic possibilities and limits of this attractive graphic form, which is very often used in the Internet for displaying various data.

Keywords: Circular visualization, Chord diagram, Commuting data, Circos, D3

Úvod do problematiky dátovej vizualizácie

Súčasná informačná spoločnosť je identifikovaná využívaním rôznorodých informačných služieb v online prostredí, ako aj generovaním obrovského množstva dát. Zdieľanie týchto dát, či už v prostredí sociálnych sietí (fotografie, videá) alebo v špecifických dátových úložiskách, čelí v súčasnosti mnohým výzvam. Jednou z dôležitých oblastí zaoberajúcou sa efektívnou distribúciou týchto dát koncovým užívateľom je oblasť dátovej vizualizácie. Existuje značný počet vizualizačných metód a techník, ktoré dokážeme aplikovať na dáta rozličného charakteru. Cieľom nášho príspevku je poukázať na možnosti cirkulárnej vizualizácie dát so zameraním na podskupinu tzv. chord diagramov. Práve rôzne formy cirkulárnej vizualizácie sa v ostatnom období výrazne rozšírili pri tvorbe rozličných grafických výstupov či už v odborných vedeckých článkoch, ale aj v komerčných médiách, ktoré prezentujú aktuálne javy, udalosti v podobe variabilných infografík a prezentačných plagátov. Kým prejdeme k samotnej podstate nášho príspevku, a teda k predstaveniu samotného konceptu cirkulárnej vizualizácie, je dôležité uviesť si samotný význam dátovej vizualizácie a čo tento proces predstavuje.

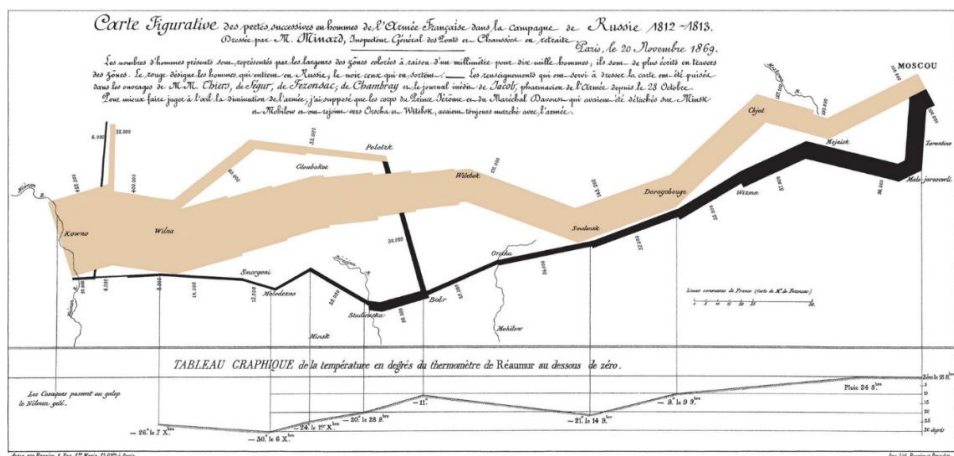
Zmyslom vizualizácie je zobrazenie čísiel, resp. dát, z ktorých na základe ich vizuálneho zobrazenia vieme získať informáciu, ktorá je v inej podobe skrytá. Najjednoduchším príkladom môže byť znázornenie časového radu vývoja počtu obyvateľov v ľubovoľnej priestorovej jednotke. Pri klasickom tabuľkovom znázornení nemôžeme pozorovať základný trend (nárast, pokles, stagnácia), ktorý je badateľný v ich grafickom vyjadrení. Základným faktom je skutočnosť, že ľudský mozog spracováva oveľa efektívnejšie informácie v ich vizuálnej podobe.

Vizualizácia dát umožňuje cieľovej skupine rýchlo získať názor na dátový rozsah (Wang a kol. 2008). Rôznorodé techniky dátovej vizualizácie pomáhajú prezentovať veľké množstvo

informácií v čo najefektívnejšej vizuálnej podobe (Kirk 2012). Samotný výstup umožňuje z komplikovaných dát vytvoriť obraz, v ktorom je na prvý pohľad viditeľné aj relatívne skryté (Ali a kol. 2016). Pri tvorbe vizualizácie je veľmi dôležitý výber metódy, ktorá je vhodná pre dané údaje a ich znázornenie. Často sa stretávame s použitím nesprávneho typu grafu, čo výslednú interpretáciu devalvuje. Vizualizácia je správna, ak si po jej preštudovaní čitateľ položí otázku, týkajúcu sa predmetných dát, resp. možných variantov ich explanácie. Vizuálne zobrazené údaje sú ľahšie pochopiteľné a analyzovateľné. Pre decízorov je dôležitá aj samotná rýchlosť nájdenia vzorcov a porozumenie zložitým vzťahom medzi entitami, ktoré dané dáta prezentujú. Vizualizácia údajov tak môže byť nápomocná pri identifikácii problémov a deficitov, pri výbere najlepšej stratégie fungovania produktu a firmy, pri prognózovaní objemu predaja a cien akcií, dolaďovania riadenia projektov, optimalizácii dopravných tokov, atď. (Bylinskii a kol. 2017). Existuje množstvo základných motívov pre využitie vizualizácie dát. Medzi tieto možno zaradiť aj (upravené podľa Pursel 2005):

- Vysvetlenie údajov alebo uvedenie údajov do kontextu (zvýraznenie podstatných vecí)
- Identifikácia problémových oblastí (rôzne extrémne výkyvy, abnormality)
- Identifikácia časových, priestorových a objemových špecifik
- Zvýraznenie existujúcich väzieb medzi skúmanými entitami
- Zvýraznenie alebo ilustrovanie inak neviditeľných údajov (napríklad izolovať odľahlé údaje, ktoré sa nachádzajú v údajoch).

Hybnou silou dátovej vizualizácie bol vývoj počítačovej technológie. Počítače umožnili spracovanie veľkého množstva údajov veľkou rýchlosťou. Dnes sa vizualizácia údajov stala rýchlo sa rozvíjajúcou zmesou vedy a umenia a v súčasnosti je vizualizácia údajov mnohými vedeckými disciplínami považovaná za moderný ekvivalent vizuálnej komunikácie (Pursel 2015). Dátová vizualizácia výrazne zmenila pohľad na dáta a ich interpretáciu. V súčasnosti si s pojmom dátovej vizualizácie spájame prezentáciu zvyčajne dynamických, interaktívnych objektov, ktoré užívateľsky prijateľným spôsobom komunikujú s klientom a sprostredkovávajú mu nejakú informáciu, ktorú chceme zdôrazniť. Počítače jednoznačne akcelerovali vývoj v tejto oblasti, avšak samotná problematika vizualizácie dát siaha do dávnej minulosti. Z historického pohľadu je zaujímavá štúdia Friendly (2008), v ktorej autori poukázali na historické míľniky vo využívaní štatistických, kartografických a grafických metód v oblasti dátovej vizualizácie. Za jedno z prelomových multivariačných diel dátovej vizualizácie je považovaná práca Charlesa Minarda z roku 1812, v ktorej zmapoval Napoleonovu inváziu do Ruska, pričom jeho grafický výstup kombinoval viaceré parametre (počet vojakov, teplota, vzdialenosť, smery, geografické súradnice, čas), z dôvodu lepšieho pochopenia kauzality tejto udalosti (Friendly 2008), (obr. 1). Podľa E. Tufte (2001) ide o najlepší príklad grafiky, aká bola kedy vytvorená.



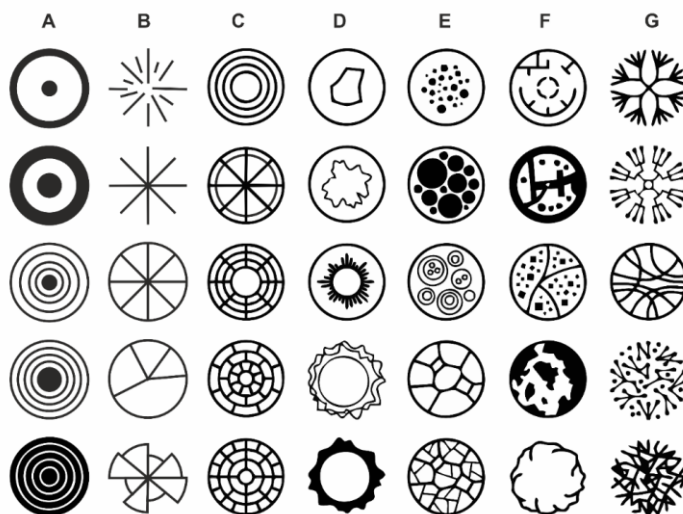
Obr. 1. Práca Charlesa Minarda z roku 1812 mapujúca Napoleonovu inváziu do Ruska

Zdroj: Grace (2018)

Dátová vizualizácia je veľmi dôležitou súčasťou vedeckej práce. V niektorých oblastiach sú vizuálne výstupy absolútnou prioritou, v iných predstavujú podporné mechanizmy pri zhodnotení získaných výsledkov. Každopádne treba konštatovať, že súčasný trend na vizuálne príťažlivú prezentáciu (nielen geografických a štatistických dát) je pomerne výrazný, a je preto dôležité venovať pozornosť tejto oblasti. Často je práve nevhodne zvolená forma vizualizácie, nedodržanie základných pravidiel jej tvorby kľúčovým faktorom pri hodnotení úspechu, resp. neúspechu či už vedecky korektného výstupu, alebo z pohľadu „online“ sveta (ne)návštevnosti internetovej stránky prezentujúcej svoj obsah zameraný práve na efektívne zobrazenie dát, resp. výsledkov vedeckého výskumu.

Cirkulárna vizualizácia

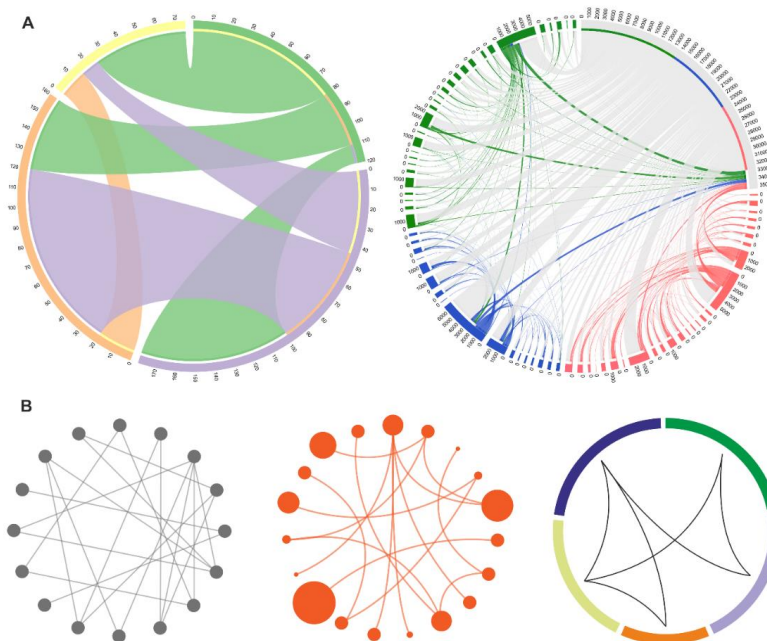
V ostatných rokoch sa cirkulárny layout grafického spracovania dát stal pomerne dominantným vo svete dátovej vizualizácie. Dôvody, prečo je tomu tak, prináša príspevok od Lima (2017), ktorý poukazuje na viaceré skutočnosti tejto preferencie z pohľadu vedeckých štúdií a výskumov zaoberajúcich sa vnímaním tvarov a objektov ľuďmi. Existuje všeobecná téza prirodzenej preferencie vyhladených kriviek nad ostrými črtami. Ľudský organizmus totiž dáva prednosť tvarom a predmetom, ktoré evokujú bezpečnosť, čo je charakteristické práve pre vyhladené povrchy a krivky. Naopak predmety s ostrými uhlami a špicovými črtami naznačujú hrozbu a zranenie. Krivka, ktorá je konečným zakriveným tvarom, predstavuje všetky atribúty, ktoré nás priťahujú: je to bezpečný, jemný, príjemný, elegantný tvar, ktorý evokuje pokoj a relaxáciu. Tieto tvrdenia sú podchytené viacerými štúdiami. Samotná náklonnosť pre estetickú preferenciu kruhu je teda zakorenená v základoch biológie ako takej. Pri cirkulárnej vizualizácii máme k dispozícii široké spektrum základných vyjadrení, ktoré možno uplatniť pri znázornení špecifických javov (obr. 2).



Obr. 2. Príklady cirkulárnej vizualizácie dát (spracované podľa Lima 2017)
 Vysvetlivky: A – kruhy a špirály | B – výseky | C – mriežky a rastre | D – prírastky,
 úbytky, toky | E – tvary a hranice | F – mapy a plány | G – uzly a spojenia

V našom príspevku sa zameriame na konkrétnu skupinu a to tzv. “chord diagram”. Na úvod je potrebné venovať pozornosť základnej terminológii súvisiacej s týmto označením grafickej interpretácie. Tento termín je bežne používaný v anglicky písanej literatúre (Abel, Sander 2014, Andersen et al. 2016, Flajolet, Noy 2000). Tvorcovia rôznych digitálnych knižníc a programov na vizualizáciu taktiež používajú označenie ako „Radial network diagram”, „Chord layout” (D3) alebo „Dependency wheel” (Highcharts.js). Vychádzajúc z jeho podstaty by bolo

možné použiť označenie ako cirkulárny diagram, toto pomenovanie sa však hodí na oveľa väčšie spektrum grafov, ktoré zodpovedajú takémuto označeniu. Naznačeniu existencie väzieb medzi jednotlivými entitami v tomto grafe by mohlo zodpovedať slovo „súvislosť alebo vzťah“. Z tohto by sa dalo teda odvodiť označenie ako „Cirkulárne vyjadrenie súvislosti medzi entitami (uzlami)“ alebo „Cirkulárny diagram súvislosti“, tu sa však dostávame do pomerne zložitého označenia, a vzhľadom na jasné označenie grafu v anglickej literatúre budeme aj v našom príspevku preferovať označenie „chord diagram“. Samotná podstata vizualizácie spočíva v znázornení vzájomnej súvislosti medzi jednotlivými entitami (priestorové jednotky, osoby, chemické prvky, atď.), ktorých rozloženie je aranžované v kruhovom zoskupení. Samotné spojenia medzi entitami sú veľkosťne prispôbené významu daného spojenia (toku). Existuje aj tzv. „non-ribbon chord diagram“ (bezpruhový chord diagram), ktorý je akousi zjednodušenou verziou chord diagramu, kedy je dôraz položený na znázornenie existencie spojení medzi entitami, ako na ich kvantifikáciu. Samotná kvantifikácia v takomto prípade môže byť súčasťou (aj nemusí) grafickej interpretácie symbolu konkrétnej entity (uzla) (obr. 3.).

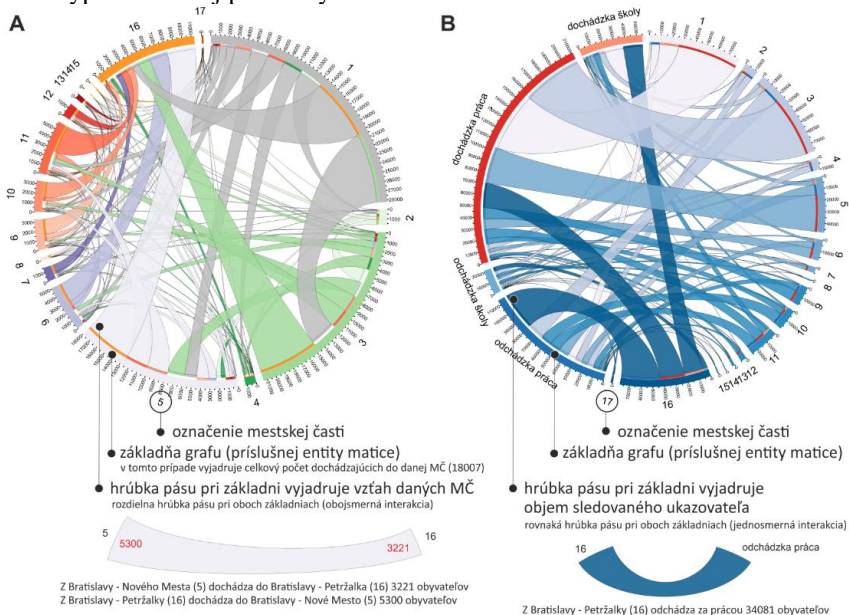


Obr. 3. Porovnanie chord diagramu a bezpruhového chord diagramu (non-ribbon)
Vysvetlivky: A – chord diagram | B – „non-ribbon chord diagram“

Vizuálna podoba výsledného diagramu môže na prvý pohľad pôsobiť veľmi zložito, preto treba vysvetleniu jednotlivých elementov venovať náležitú pozornosť (viď časť „Dáta a použité nástroje“). Vzhľadom na túto skutočnosť sa ako vhodné javí poukázať na základnú formu interpretácie takéhoto grafu. Základný spôsob čítania grafu si ukážeme na dvoch príkladoch: Diagram so vzájomným vzťahom medzi entitami (dochádzka za prácou medzi jednotlivými mestskými časťami) a diagram bez vzájomného vzťahu medzi entitami (Dochádzka a odchádzka za prácou a do škôl z jednotlivých mestských častí Bratislavy) (obr. 4). Detailnejšie rozdiely v zápise vstupných dát potrebných pre konštrukciu takýchto diagramov sú rozpísané v ďalšej časti príspevku.

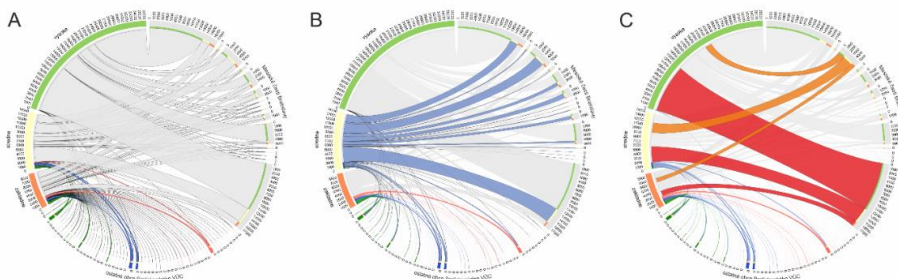
O čitateľnosti výsledného diagramu rozhoduje viacero faktorov (poradie entít, farebnosť spojení, počet prechodov oblúkom, transparentnosť, atď.), ktoré môže ovplyvniť priamo tvorca grafiky. Z pohľadu samotnej realizácie vykreslenia diagramu ide o aplikovanie matematických operácií, ktoré sú integrované v konkrétnej digitálnej knižnici. Niektoré z postupov konštrukcie spojení sú opísané v prácach (Stoimenow 2000, Phan a kol. 2005, Andersen a kol. 2016).

Jedným z problémov vykreslenia spojení medzi uzlami je ich veľké množstvo, ktoré môže byť súčasťou vstupných dát. Na zníženie vizuálnej zložitosti takéhoto znázornenia sa používa technika zvaná hierarchické zoskupovanie hrán (Holten 2006, Zhou a kol. 2013). Jednou z možností ako eliminovať veľké množstvo spojení medzi entitami je upriamiť pozornosť čitateľa na podstatné informácie z diagramu, resp. na tie, ktoré považujeme z hľadiska interpretácie za kľúčové. Toto možno dosiahnuť primárnou editáciou parametrov diagramu a jeho výslednou úpravou (obr. 5). V prípade online riešenia máme výraznejšie možnosti identifikovať entity s veľmi nízkymi hodnotami. Tu je vhodné využiť interaktivitu, animácie výsledného grafu, eventuálne vypnutie menej podstatných uzlov.



Obr. 4. Príklad čítania chord diagramu

Vysvetlivky: A – vzájomný vzťah medzi entitami | B – bez vzájomného vzťahu medzi entitami
Pôvodný obrázok vo väčšom rozlíšení uvádza Bačík (2019a)

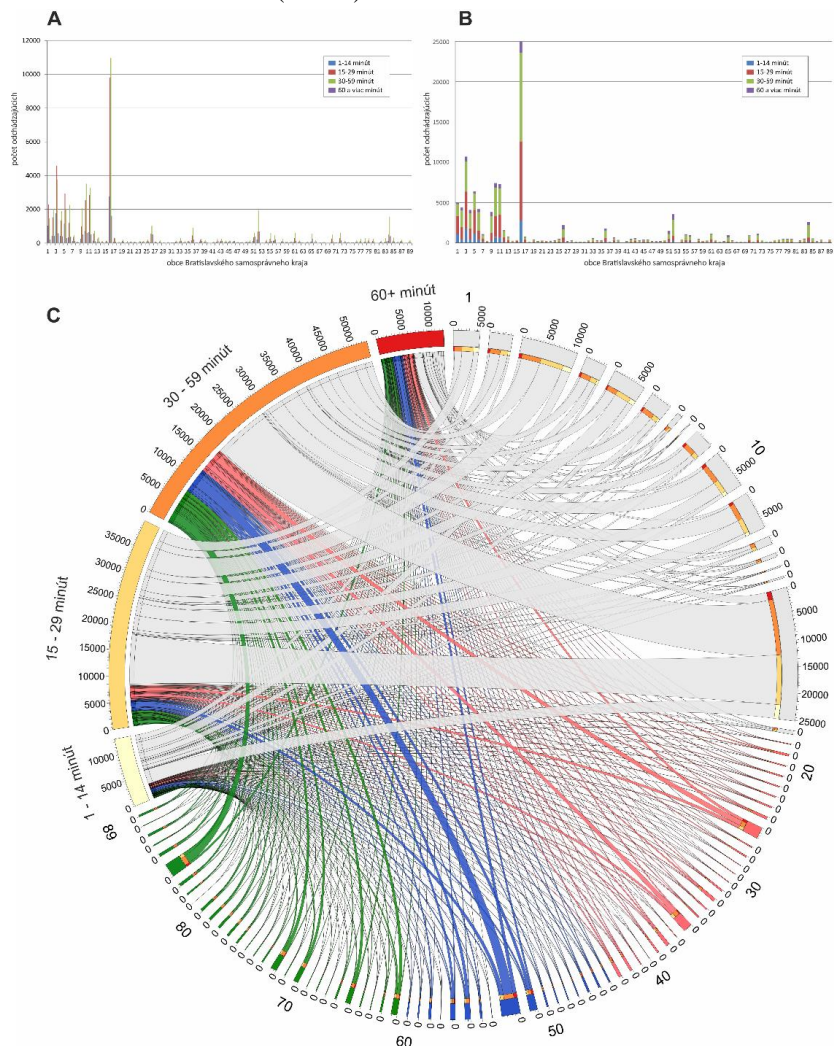


Obr. 5. Rôzne preferencie zobrazenia dát v chord diagrame po úprave

Vysvetlivky: A – primárny grafický výstup bez editácie elementov | B, C – diagram po editácii, so zacielením na konkrétne dáta

Prímárne sa použitie chord diagramov uplatňuje pri vizualizácii objektov, medzi ktorými existuje priamy vzťah. V demografii využili tento koncept ako prví Abel, Sander (2014), Sander a kol. (2014), ktorí znázornili týmto grafickým vyjadrením svetovú migráciu. Práve migračné a dochádzkové dáta sú veľmi vhodné na použitie chord diagramov, nakoľko dobre ilustrujú vzájomné vzťahy medzi destinačnými a cieľovými oblasťami. Chord diagram je však

možné použiť aj v prípadoch kedy medzi entitami neexistuje vzájomná interakcia a ide v podstate o jednoduchú kvantifikáciu dát (podobne ako v stĺpcovom grafe). Jeden súbor entít môže predstavovať konkrétne jednotky (napr. regióny) a ďalšie entity kategórie sledované v grafe (viď. Obr. 4B). Ako príklad použijeme znázornenie odchádzky za prácou podľa dĺžky trvania v obciach Bratislavského VÚC¹ (obr. 6.).



Obr. 6. Porovnanie stĺpcového, skladaného grafu a chord diagramu zobrazujúceho odchádzku podľa dĺžky trvania (Šírka základne diagramu pri jednotlivých obciach predstavuje celkový počet denne odchádzajúcich za prácou z danej obce a šírka základne pri časových intervaloch vyjadruje celkový objem odchádzajúcich v danom časovom intervale z obcí Bratislavského VÚC); Vysvetlivky: A – Stĺpcový graf (os x – obce Bratislavského VÚC, os y – počet odchádzajúcich v danom časovom intervale) | B – Skladaný graf (os x – obce Bratislavského VÚC, os y – počet odchádzajúcich v danom časovom intervale) | C – Chord diagram (v smere hodinových ručičiek: Čísla 1 – 89 označujú obce Bratislavského VÚC, ďalej nasledujú počty odchádzajúcich v daných časových intervaloch);

¹ VÚC – Vyšší územný celok. V príspevku používame aj označenie samosprávny kraj, ktorý z pohľadu organizácie verejnej správy vyjadruje tú istú územnú jednotku.

V tomto prípade sú spojenia znázornené len medzi danou obcou a konkrétnymi časovými kategóriami (nie medzi obcami navzájom). Pri použití klasického stĺpcového, resp. skladaného grafu, môžeme identifikovať objem odchádzky za konkrétny čas pre konkrétne obce, avšak pri použití chord diagramu vieme okamžite identifikovať okrem parciálnych počtov aj celkový objem odchádzajúcich v konkrétnom časovom intervale. V uvedenom grafe je to skutočnosť, že väčšina žiakov trvá cesta do školy v rozpätí 30 – 59 minút. Pri použití štandardného grafického vyjadrenia ostáva tento údaj viac-menej ukrytý, resp. identifikácia tohto objemu je pomerne náročná. Rovnako tak, ako vidieť z obrázku, možno konštatovať, že pri väčšom počte entít je vhodnejšie používať náhradu štandardných grafov, nakoľko počet zobrazovaných jednotiek na osi x (obce Bratislavského VÚC) je z hľadiska jednoduchosti interpretácie pomerne limitujúci.

Na záver krátka poznámka k histórii používania chord diagramov. V úvode príspevku sme spomenuli, že samotná problematika dátovej vizualizácie siaha niekoľko storočí do minulosti. Z tohoto pohľadu je nami sledovaný chord diagram skutočným „nováčikom“ medzi rozličnými grafikami. Prvý chord diagram bol publikovaný v roku 2007 v New York Times a zobrazoval detailné informácie o genóme. Od tohto obdobia však jeho využiteľnosť výrazne narástla a stala sa súčasťou vizualizácie javov takmer vo všetkých vedných disciplínach opisujúcich rozličné formy vzťahov medzi objektami, geografiu nevyímajúc. Popularita takéhoto grafického vyjadrenia vyústila aj do priamej implementácie výpočtov potrebných pre zobrazenie chord diagramov do rozličných programov a vizualizačných knižníc. Z mnohých nástrojov možno menovať napríklad Flourish, amCharts, Tableau, R, ZingChart, Highcharts, D3 a Circos. Práve posledné dva menované produkty sme použili aj v našej vizualizácii dochádzkových vzťahov na príklade obcí Bratislavského VÚC.

Dáta a použité nástroje

Pri tvorbe nášho príspevku sme využili dáta o dochádzke zo Sčítania obyvateľov, domov a bytov z roku 2011. Sme si vedomí, že práve niektoré dáta týkajúce sa dennej dochádzky do práce boli pri tomto sčítaní podrobené pomerne výraznej kritike, avšak pre naše účely a vzhľadom na použitú mierku ich spracovania (obce BA VÚC) boli postačujúce. Podoba originálneho vstupného súboru je na obr. 7.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	TAB. 702 Bývajúce obyvateľstvo odchádzajúce do zamestnania a škôl podľa pohlavia a veku a podľa obce odchádzky a obce dochádzky														
2	SODB 2011														
3	Okres Bratislava														
4	Obec odchádzky	Obyvateľstvo ekonomicky aktívne odchádzajúce do zamestnania										Žiaci a študenti			Odchádzka
5	obec dochádzky	muži	ženy	spolu	v tom vo veku					spolu	z toho			spolu	
6					15-24	25-29	30-34	35-44	45-59	60+		žiaci ZŠ	študenti SS	študenti VŠ	
7	a	1	2	3	4	5	6	7	8	9	10	11	12	13	14
8	OKRES HLCHOVEC														
9	Uhň odchádzajúcich v okrese	5 371	4 056	9 427	772	1 457	1 473	2 674	2 863	188	2 795	594	1 122	1 079	12 222
10	v tom odchádzajúci v rámci okresu	1 886	1 610	3 496	238	352	491	1 109	1 228	78	734	433	273	28	4 230
11	odchádzajúci do iných okresov SR	3 211	2 206	5 417	498	981	822	1 442	1 569	105	1 972	153	840	979	7 389
12	z toho okres:														
13	Bratislava hl m. SR	783	456	1 239	96	337	218	305	264	19	342	7	22	313	1 581
14	Košice - mesto	1	2	3	0	0	1	2	0	0	4	0	0	4	7
15	Bánovce nad Bebravou	3	2	5	1	0	1	1	2	0	2	2	0	0	7
16	Banská Bystrica	6	1	7	2	0	2	1	1	1	20	0	0	20	27
17	Brezno	2	0	2	0	1	1	0	0	0	0	0	0	0	2
18	Čadca	1	0	1	0	0	1	0	0	0	2	0	1	1	3
19	Dunajská Streda	5	0	5	0	1	0	1	3	0	1	0	1	0	6
20	Galanta	137	116	253	25	40	31	60	92	5	32	4	21	7	285
21	Iľava	4	0	4	0	0	2	0	1	1	0	0	0	0	4
22	Komárno	0	1	1	0	0	0	1	0	0	0	0	0	0	1
23	Levice	23	3	26	1	5	2	4	13	1	1	1	0	1	27
24	Levoča	1	1	2	0	0	0	0	2	0	0	0	0	0	2
25	Liptovský Mikuláš	0	1	1	0	0	0	0	1	0	0	0	0	0	1
26	Malacky	19	4	23	2	4	8	5	4	0	2	2	0	0	25
27	Modra	3	1	4	0	1	0	0	2	1	1	0	0	1	5
28	Martin	1	0	1	0	0	0	0	1	0	0	0	0	0	1
29	Michalovce	5	1	6	0	0	1	3	2	0	1	0	0	1	7
30	Myjava	188	188	376	41	67	55	101	107	5	320	2	67	251	696
31	Nitra	29	34	63	7	9	13	16	18	0	6	2	4	0	69
32	Nové Zámky	8	0	8	0	1	3	2	2	0	0	0	0	0	8
33	Partizánske	1	0	1	0	0	1	0	0	0	0	0	0	0	1
34	Pezinok	28	8	36	5	8	4	12	7	0	14	0	10	4	50
35	Piešťany	310	344	654	59	104	88	177	214	12	268	30	217	21	922
36	Považská Bystrica	6	0	6	0	3	1	1	1	0	0	0	0	0	6
37	Prešov	3	0	3	0	0	0	2	1	0	1	1	0	0	4
38	Príevazda	2	1	3	0	0	0	1	2	0	0	0	0	0	3
39	Púchov	0	0	0	0	0	0	0	0	0	1	0	0	1	1
40	Rimavská Sobota	3	2	5	5	0	0	0	0	0	0	0	0	0	5
41	Ružomberok	1	1	2	1	1	0	0	0	0	3	0	0	3	5
42	Senec	34	30	64	4	12	9	18	20	1	3	0	3	0	67
43	Senica	7	2	9	1	2	3	0	3	0	4	0	3	1	13
44	Šalica	1	2	3	0	2	1	0	0	0	2	0	0	2	5
45	Šaľa	4	3	7	0	2	1	0	3	1	1	0	1	0	8

Obr. 7. Ukážkový hárok vstupného súboru o dochádzke

Samotný súbor obsahuje 4 hárky (2 pre údaje o dochádzke a 2 pre údaje o odchádzke). Štruktúra jednotlivých hárkov nám vyhovovala z hľadiska sledovaných atribútov o dochádzke a odchádzke, ktoré sme následne vizualizovali pomocou chord diagramov. Pri oboch procesoch vieme z týchto dát zistiť množstvo ukazovateľov: počet mužov a žien podieľajúcich sa na dochádzke, ich vekovú štruktúru, čas dochádzky v minútach ako aj odvetvie hospodárstva (pri dochádzke podľa ekonomickej činnosti), resp. typ školy (pri hodnotení dochádzky do škôl). Okrem toho samozrejme nájdeme smery dochádzky pre každú jednotlivú obec. Vzhľadom na skutočnosť, že tieto dáta prezentujeme aj online v podobe internetovej stránky (http://www.sodbtn.sk/dochadzka_ba), bolo potrebné pri spracovaní dát využiť množstvo viacerých metód a postupov (priradovanie kódov obcí do atribútov, zápis do databázových tabuliek na server, SQL funkcie výpisu dát pre konkrétnu obec, prenos premenných medzi stránkami a pod.) V našom príspevku sa však zameriavame na využitie cirkulárnej vizualizácie týchto dát (pomocou chord diagramov), a preto sa nebudeme týmto dielčím postupom detailnejšie venovať. Zameriame sa na opis online rozhrania „Circos“, pomocou ktorého možno vytvárať chord diagramy a tieto výstupy ďalej editovať podľa potreby, ako aj na využitie knižnice D3, pomocou ktorej sme vytvárali interaktívne verzie chord diagramov z predmetných dát a následne distribuovali v prostredí Internetu.

Online aplikácia Circos

Circos je softvérový balík na vizualizáciu údajov a informácií, ktoré vizualizuje v kruhovom usporiadaní, čím je vzhľadom na súčasné trendy v oblasti grafickej vizualizácie veľmi atraktívny (Krzywinski a kol. 2009). Tvorcom projektu Circos je Martin Krzywinski (Genome Sciences Center) a všetky potrebné informácie možno nájsť práve na stránkach vytvorených autorom projektu (<http://circos.ca>). Pôvodne bol navrhnutý na vizualizáciu genomických údajov, avšak jeho popularita veľmi rýchlo narástla a je vhodný pre akýkoľvek odbor v oblasti vedy alebo umenia, o čom svedčí aj obrovské množstvo dostupných vizualizácií vytvorených v tomto programe, ktoré sú dostupné priamo na stránkach produktu. Je určený pre všetkých, ktorí chcú znázorniť alebo preskúmať veľké množstvo údajov. V prípade našich dát sme využili online nástroj „TableViewer“, ktorý je dostupný na stránkach projektu Circos (Krzywinski et al. 2009). Tento nástroj pracuje s dátami zapísanými v podobe matice vo formáte *.txt, pričom jednotlivé hodnoty musia byť oddelené tabulátormi. Výsledok si môžeme ukázať na modelovom príklade piatich imaginárnych regiónov (A – E), medzi ktorými existujú isté vzťahy. Môžeme si pod týmito prvkami predstaviť jednotlivé obce, medzi ktorými vzájomne navzájomné dochádzkové väzby. Štruktúra vstupných dát je zobrazená v tab. 1.

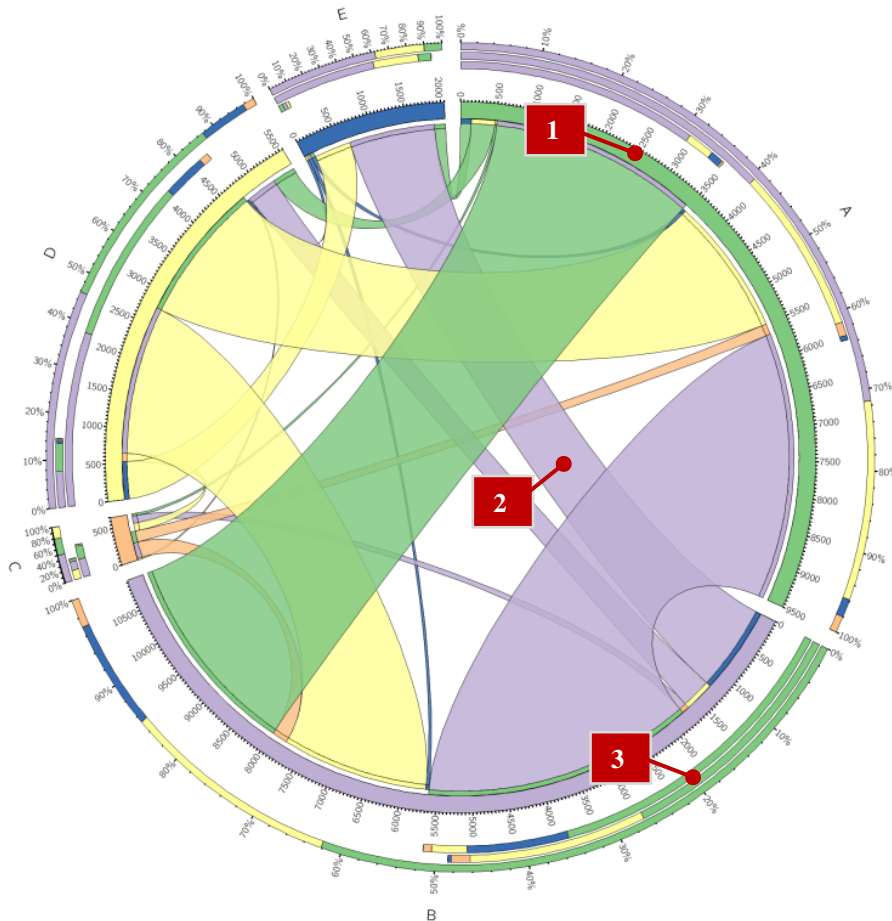
Tab. 1. Zápis dát potrebných pre ich cirkulárnu vizualizáciu v prostredí CIRCOS

data	data	data	6	7	8	9	10
data	data	data	127,201,127	190,174,212	253,192,134	255,255,153	56,108,176
data	data	data	A	B	C	D	E
1	127,201,127	A	0	2940	32	332	154
2	190,174,212	B	3852	0	103	427	1234
3	253,192,134	C	159	223	0	19	0
4	255,255,153	D	1979	2111	115	0	531
5	56,108,176	E	53	44	0	36	0

	označenie poradia v akom sa vykresľujú jednotlivé segmenty
	farebné označenie segmentov (RGB zápis)
	jednotlivé segmenty matice

Samotná dáta nemusia obsahovať stĺpce označenia poradia v akom sa vykresľujú, ani hodnoty farieb príslušných elementov v podobe zápisu príslušného RGB kódu. Pri ich absencii je farba a poradie doplnená automaticky. Z užívateľského hľadiska je však vhodnejšie tieto hodnoty zahrnúť do vstupnej matice. V prípade nekorektného zápisu dát sa zobrazia namiesto

výsledného chord diagramu hlásenie, ktoré nás upozorní na chybu v dátach. Pri korektnom zápise vstupných hodnôt je výsledkom základná podoba chord diagramu, ktorého jednotlivé položky sú vysvetlené na obr. 8.

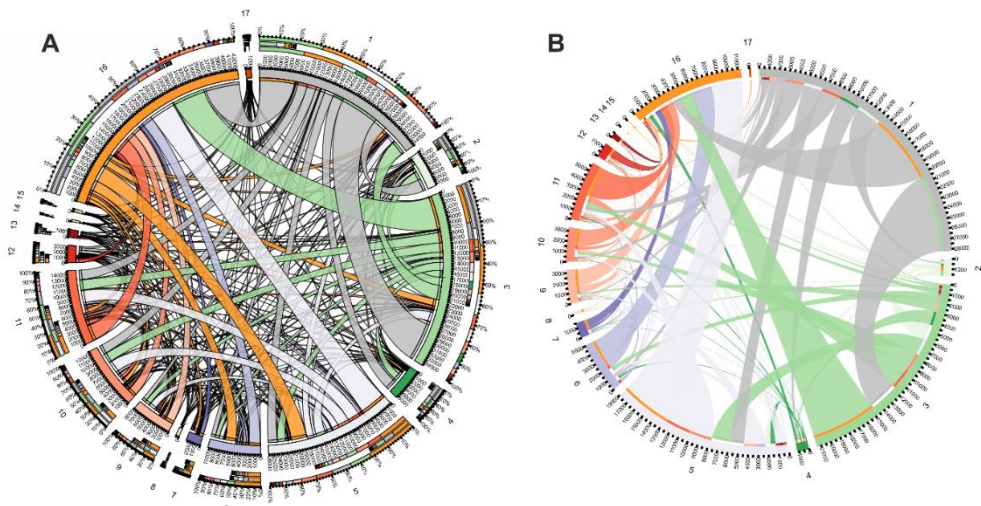


Obr. 8. Základná podoba chord diagramu generovaného službou CIRCOS pri správnom nastavení vstupných hodnôt

Vysvetlivky: 1 – Segmenty riadkov a stĺpcov matice; 2 – Šírka pruhu (ribbon) zobrazuje hodnotu bunky spojenej s párom segmentu riadok/stĺpec; 3 – Relatívne vyjadrenie riadku, stĺpca a celkovo pre každý segment

Primárne nastavenie parametrov vykreslenia však vo väčšine prípadov nevytvorí diagram, ktorý by bol vhodný pre zobrazenie nami požadovaných informácií. A preto je súčasťou tohto online produktu aj užívateľské rozhranie, v ktorom môže tvorca diagramov nastaviť preferenciu výsledného diagramu. Takéto porovnanie medzi základným nastavením upravenými parametrami diagramu možno vidieť na obr. 9, v ktorom je znázornená dochádzka za prácou medzi jednotlivými mestskými časťami Bratislavy.

Výsledná podoba diagramu teda závisí od príslušných nastavení, ktoré sú dostupné pri využití online služby. Okrem toho je možné nainštalovať aj samotný program CIRCOS, ktorý ponúka ešte výraznejšie možnosti výslednej vizualizácie. Samotný výsledný výstup je generovaný ako obrázok vo formáte PNG, ale zároveň aj vo vektorovom formáte SVG. Tento možno načítať v akomkoľvek vektorovom grafickom programe, čo nám dáva veľký priestor pre ďalšiu editáciu a modifikáciu diagramu tak, aby zodpovedal tomu, čo chceme primárne danou vizualizáciou ukázať.

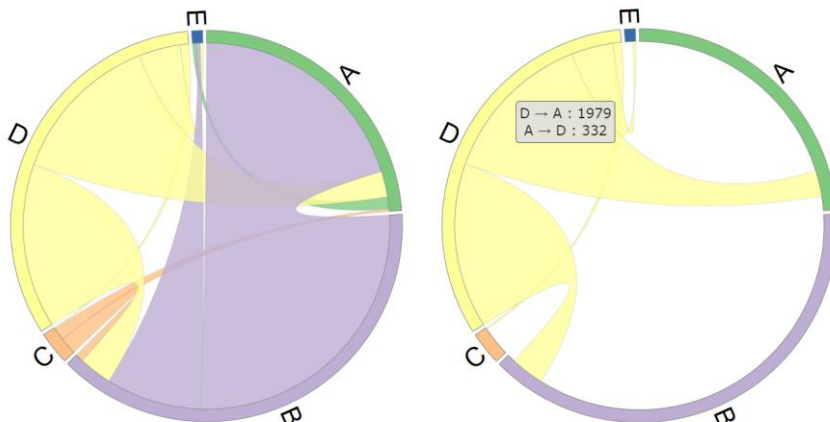


Obr. 9. Rôzne nastavenie atribútov chord diagramu vytvoreného pomocou Circos-u
 Vysvetlivky: A – originálne nastavenie zobrazenia | B – modifikácia vstupných parametrov (skrytie relatívneho vyjadrenia, usporiadanie pruhov pre jednotlivé segmenty vzostupne, vyfarbenie pruhov podľa riadku v matici, odstránené orámovanie elementov grafu a racionalizované vykreslenie pruhov; Názvy mestských častí (v smere hodinových ručičiek): 1 – Bratislava-Staré Mesto | 2 – Bratislava-Podunajské Biskupice | 3 – Bratislava-Ružinov | 4 – Bratislava-Vrakuňa | 5 – Bratislava-Nové Mesto | 6 – Bratislava-Rača | 7 – Bratislava-Vajnory | 8 – Bratislava-Devín | 9 – Bratislava-Devínska Nová Ves | 10 – Bratislava-Dúbravka | 11 – Bratislava-Karlova Ves | 12 – Bratislava-Lamač | 13 – Bratislava-Záhorská Bystrica | 14 – Bratislava-Čunovo | 15 – Bratislava-Jarovce | 16 – Bratislava-Petržalka | 17 – Bratislava-Rusovce

D3 knižnica

Pri výbere nástroja vhodného pre interaktívnu verziu chord diagramov zohrávala úlohu samotná znalosť prostredia, v ktorom sa daný diagram vytvára. Ako už bolo spomínané v úvode, dynamické chord diagramy možno vytvárať pomocou viacerých nástrojov. Digitálna knižnica D3 už istý čas predstavuje dynamický prístup k vizualizácii geografických údajov (aj negeografických, resp. nepriestorových dát) publikovaných cez webové rozhranie. Pomocou D3 môžu vývojári spájať vstupné dáta do ľubovoľných prvkov dokumentu, aplikovať dynamickú transformáciu a modifikovať ich obsah (Bostock, Ogievetsky, Heer, 2011). Chord diagramy vytvorené pomocou tejto knižnice sú na rozdiel od diagramov vytvorených pomocou Circos-u, interaktívne a umožňujú tak detailné zobrazenie akejkoľvek informácie v grafe. Aplikácia tejto knižnice je výhodná práve vzhľadom na možnosti priamej interakcie klienta s diagramom. Dôležitým momentom pri výbere D3 knižnice bola aj skutočnosť, že zdrojové dáta môžu byť zapísané vo forme matice (s modifikáciou zápisu kompatibilného s príslušnými funkciami knižnice), práve tak, ako tomu bolo pri vstupných súboroch potrebných pre vytvorenie diagramov pomocou Circos-u. Príklad si ukážeme na rovnakých dátach, ako tomu bolo v prípade chord diagramu vytváraného pomocou Circos-u (ide o dáta z tab. 1) (obr. 10).

Výsledný diagram automaticky využíva racionalizované vykreslenie pruhov, t. z., že hrúbka pruhu na jednom konci je iná ako na druhom. Inými slovami povedané, hrúbka pásu na oboch koncoch zodpovedá reálnym hodnotám sledovaného ukazovateľa v smere „z“ „do“. Farebné znázornenie jednotlivých segmentov a pruhov je napísané vo vstupnom súbore v hexadecimálnom zápise. Hlavnou výhodou je interaktivita diagramu, kedy možno zobraziť presné dielčie hodnoty jednotlivých pruhov, ako aj segmentov vstupnej matice.



Obr. 10. Chord diagram vytvorený pomocou knižnice D3

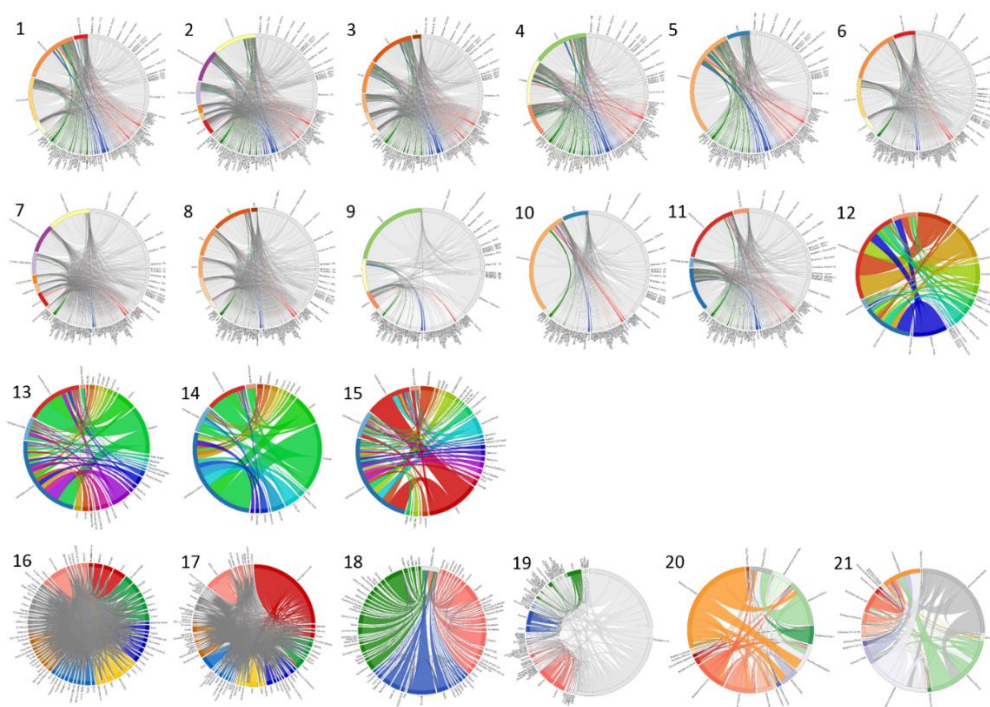
Výsledky

Naším cieľom bolo vytvorenie série grafov (chord diagramov), ktoré by prezentovali vybrané atribúty dát o dochádzke za prácou a do škôl medzi jednotlivými obcami Bratislavského VÚC. Keďže primárne sme sa snažili o interpretáciu týchto dát v online prostredí, dominovalo využitie knižnice D3 a jej funkcií pred diagramami vytváranými v prostredí Circos (príklady uvádza Bačík 2019b). Tento je oveľa vhodnejšie použiť pri printovej prezentácii dát v podobe rôznych posterov a infografík, s možnosťou editácie diagramu tak, aby jeho výpovedná hodnota bola zacielená na kľúčové prvky zobrazované v týchto diagramoch (dôraz na vybrané charakteristiky, ktoré autor považuje za relevantné z hľadiska získaných výsledkov). Celkovo bolo na základe vstupných dát vytvorených 21 chord diagramov, ktoré možno rozdeliť do dvoch skupín – so vzájomnou interakciou vstupných entít a bez vzájomnej interakcie vstupných entít (tab. 2, obr. 11).

Tab. 2. Zoznam chord diagramov dostupných na stránke www.sodbtm.sk/dochadzka_ba (Bačík 2019c)

Bez vzájomnej interakcie	1	Denná odchádzka z obcí BA kraja podľa času trvania
	2	Odchádzka z obcí BA kraja podľa odvetvia hospodárstva
	3	Odchádzka z obcí BA kraja podľa veku
	4	Odchádzka z obcí BA kraja do škôl
	5	Odchádzka z obcí BA kraja do zamestnania a do škôl
	6	Denná dochádzka do obcí BA kraja podľa času trvania
	7	Dochádzka do obcí BA kraja podľa odvetvia hospodárstva
	8	Dochádzka do obcí BA kraja podľa veku
	9	Dochádzka do obcí BA kraja do škôl
	10	Dochádzka do obcí BA kraja do zamestnania a do škôl
	11	Dochádzajúci a odchádzajúci do práce a do škôl v obciach BA kraja
	12	Dochádzajúci a odchádzajúci do práce a do škôl v mestských častiach Bratislavy
	13	Dochádzajúci a odchádzajúci do práce a do škôl v obciach okresu Malacky
	14	Dochádzajúci a odchádzajúci do práce a do škôl v obciach okresu Pezinok
	15	Dochádzajúci a odchádzajúci do práce a do škôl v obciach okresu Senec
So vzájomnou interakciou	16	Odchádzka za prácou v okresoch SR (2011) (do iných okresov)
	17	Dochádzka za prácou v okresoch SR (2011) (z iných okresov)
	18	Odchádzajúci v obciach BA kraja medzi sebou
	19	Dochádzajúci v obciach BA kraja medzi sebou
	20	Odchádzajúci v mestských častiach BA medzi sebou
	21	Dochádzajúci v mestských častiach BA medzi sebou

V prvom prípade ide teda o diagramy, znázorňujúce dochádzkové toky z miesta A do miesta B, a zároveň z miesta B do miesta A. Hodnoty dochádzky/odchádzky sú vykreslené rôznou hrúbkou pruhu pri základni danej obce. V druhej skupine diagramov ide o znázornenie konkrétnych atribútov dochádzky (čas, vek, odvetvie hospodárstva, atď.), a teda nejde o oboj-
smernú interakciu, sledovaná je konkrétna sumárna hodnota atribútu (čas dochádzky zo všet-
kých obcí, počet žiakov základných a stredných škôl, atď.)



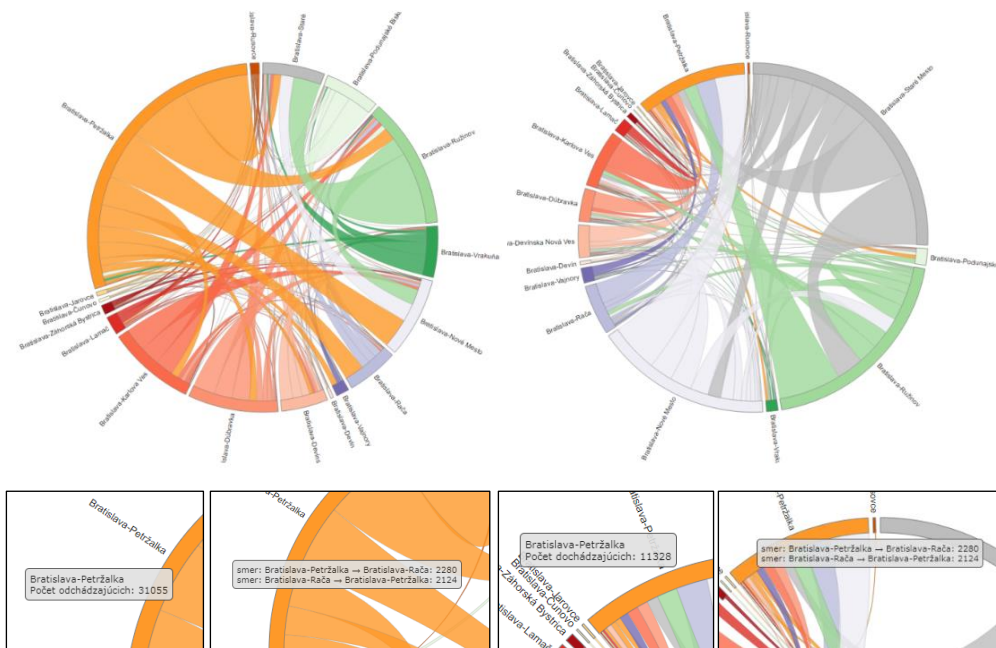
Obr. 11. Ukážky výsledných interaktívnych chord diagramov

Okrem atraktívneho kruhového usporiadania je najdôležitejším prvkom vytvorených chord diagramov interaktívnosť, možnosť získať konkrétnu hodnotu pre konkrétny dochádzkový atribút. Kompletná funkcionálnosť vykreslenia jednotlivých elementov je súčasťou použitej D3 knižnice a využitie jej výhod je možné bez akéhokoľvek zásahu do jej vnútorného obsahu. Modifikácia výsledného vzhľadu a zobrazenia dát je možná pomocou jednoduchých úprav zdrojového kódu v JavaScripte. Zobrazenie elementov a spojovacích pruhov vychádza z dvoch vstupných súborov, ktoré sú potrebné pre vykreslenie chord diagramu. Jeden súbor (polozky.csv) obsahuje jednotlivé položky matice (obce, prípadne ostatné atribúty sledovanej dochádzky). Druhý súbor obsahuje údaje matice a výsledný skript ho spracováva vo formáte *.json (matrix.json). Vždy ide o štvorcovú maticu, ktorej hodnoty sa líšia v kontexte toho či existuje alebo neexistuje vzťah medzi sledovanými entitami. Podoba oboch vstupných súborov pre diagram so vzájomným vzťahom medzi entitami je v tab. 3. V nej vidíme, že tvorba dvoch rozdielnych chord diagramov znázorňujúcich dochádzku, resp. odchádzku za prácou je výsledkom jednoduchého procesu transponovania vstupnej matice. Všetky ostatné nastavenia ostanú nezmenené a získame dva rozdielne diagramy, ktoré sa odlišujú šírkou bázy konkrétnej mestskej časti, kedy v prvom prípade táto zobrazuje celkový počet odchádzajúcich a druhá celkový počet dochádzajúcich v danej mestskej časti. Hrúbka pruhov spájajúcich jednotlivé obce má pri jej báze rôznu šírku. Táto nám vyjadruje počet dochádzajúcich v oboch smeroch. Rovnako tak na diagramoch vidíme aj sumárne počty dochádzajúcich/odchádzajúcich za jednotlivé mestské časti Bratislavy (obr. 12).

Tab. 3. Príklad zápisu vstupných súborov

<p>Súbor: polozky.csv</p> <pre> name,color Bratislava-Staré Mesto,#bdbdbd Bratislava-Podunajské Biskupice,#e5f5e0 Bratislava-Ružinov,#ald99b Bratislava-Vrakuňa,#31a354 Bratislava-Nové Mesto,#efedf5 Bratislava-Rača,#bcbddc Bratislava-Vajnory,#756bb1 Bratislava-Devín,#fee5d9 Bratislava-Devínska Nová Ves,#fcbba1 Bratislava-Dúbravka,#fc9272 Bratislava-Karlova Ves,#fb6a4a Bratislava-Lamač,#de2d26 Bratislava-Záhorská Bystrica,#a50f15 Bratislava-Čunovo,#ffffd4 Bratislava-Jarovce,#fed98e Bratislava-Petržalka,#fe9929 Bratislava-Rusovce,#cc4c02 </pre>
<p>Súbor: matrix.json (odchádzka za prácou do jednotlivých mestských častí)</p> <pre> [[0,106,2940,158,1395,238,32,9,103,145,332,46,28,0,0,770,0], [2549,0,1650,128,531,116,37,0,59,55,86,16,0,0,0,390,0], [7875,538,0,140,1929,414,103,14,155,230,427,72,33,0,0,1370,16], [2154,345,1412,0,501,160,40,0,45,55,103,24,0,0,0,373,0], [1934,97,1666,61,0,297,57,22,105,157,305,87,70,13,9,3221,42], [885,33,803,36,868,0,109,19,78,61,147,45,19,0,0,2124,27], [159,10,223,9,184,89,0,0,15,10,19,0,0,0,0,606,12], [77,0,52,0,66,17,0,0,19,16,24,0,0,0,0,28,0], [798,37,1032,83,1346,272,67,29,0,287,301,85,23,0,0,357,0], [1876,75,1930,112,2384,513,145,26,482,0,640,172,50,0,0,756,0], [1979,75,2111,139,2506,444,115,20,305,371,0,84,24,0,0,742,0], [360,0,436,33,497,98,25,0,77,102,120,0,18,0,0,147,0], [181,0,184,25,281,56,0,0,43,55,57,22,0,0,0,70,0], [79,0,50,0,29,18,0,0,10,0,16,0,0,0,0,67,23], [53,0,44,0,76,27,0,0,38,24,36,13,0,0,0,105,17], [7110,377,6762,252,5300,2280,787,213,1720,1747,3154,795,286,90,116,0,66], [140,6,124,9,114,62,18,0,67,51,88,20,0,0,18,202,0]] </pre>
<p>Súbor: matrix.json (dochádzka za prácou do jednotlivých mestských častí – transponovaná matica)</p> <pre> [[0,2549,7875,2154,1934,885,159,77,798,1876,1979,360,181,79,53,7110,140], [106,0,538,345,97,33,10,0,37,75,75,0,0,0,0,377,6], [2940,1650,0,1412,1666,803,223,52,1032,1930,2111,436,184,50,44,6762,124], [158,128,140,0,61,36,9,0,83,112,139,33,25,0,0,252,9], [1395,531,1929,501,0,868,184,66,1346,2384,2506,497,281,29,76,5300,114], [238,116,414,160,297,0,89,17,272,513,444,98,56,18,27,2280,62], [32,37,103,40,57,109,0,0,67,145,115,25,0,0,0,787,18], [9,0,14,0,22,19,0,0,29,26,20,0,0,0,0,213,0], [103,59,155,45,105,78,15,19,0,482,305,77,43,10,38,1720,67], [145,55,230,55,157,61,10,16,287,0,371,102,55,0,24,1747,51], [332,86,427,103,305,147,19,24,301,640,0,120,57,16,36,3154,88], [46,16,72,24,87,45,0,0,85,172,84,0,22,0,13,795,20], [28,0,33,0,70,19,0,0,23,50,24,18,0,0,0,286,0], [0,0,0,0,13,0,0,0,0,0,0,0,0,0,0,90,0], [0,0,0,0,9,0,0,0,0,0,0,0,0,0,0,116,18], [770,390,1370,373,3221,2124,606,28,357,756,742,147,70,67,105,0,202], [0,0,16,0,42,27,12,0,0,0,0,0,0,0,23,17,66,0]] </pre>

V zápise dát sú zvýraznené hodnoty, ktoré sú zobrazené na obr. 12



Obr. 12. Znáozornenie odchádzky a dochádzky za prácou medzi Bratislavskými mestskými časťami
 Interaktívnu verziu diagramu pre odchádzku uvádzame na stránke sobdtn.sk (Bačík 2019d).
 Interaktívnu verziu diagramu pre dochádzku uvádzame na stránke sobdtn.sk (Bačík 2019e).

Výpis príslušných hodnôt, ktoré sa zobrazujú na diagrame po prechode myšou na jednotlivé elementy (báza entity – groupTip, alebo príslušný pruh spájajúci dve entity - chordTip) je realizovaný zápisom vstupných dát v JavaScripte:

```
function groupTip (d) {
    var p = d3.format(".1%"), q = d3.format(".0f")
    return ""+d.gname+"<br/>"
    +"Počet odchádzajúcich: " + q(d.gvalue) + "<br/>";
}

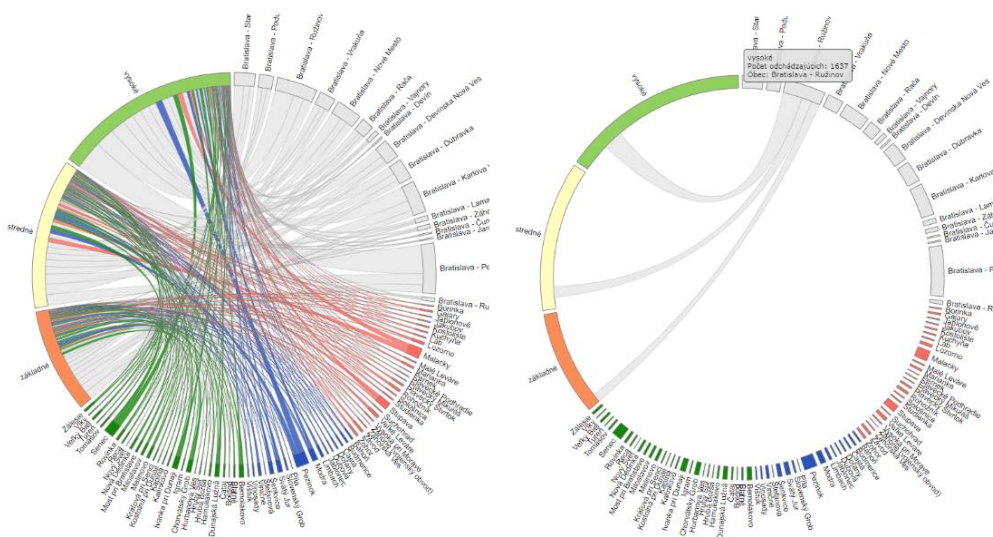
function chordTip (d) {
    var p = d3.format(".1%"), q = d3.format(",.2r")
    return ""
    + "smer: " + d.sname + " → " + d.tname
    + ": " + (d.svalue) + "<br/>"
    + "smer: " + d.tname + " → " + d.sname
    + ": " + (d.tvalue) + "<br/>";
}
```

V prípade diagramov, bez vzájomnej súvislosti je štruktúra vstupných súborov podobná, avšak pri súbore položky.csv sú na konci doplnené ďalšie sledované atribúty (napr. dochádzka do práce, dochádzka do školy, odchádzka do práce a odchádzka do školy). Maticu potom tvorí okrem údajov pre jednotlivé obce (tieto sú nulové) aj doplnenie o ďalšie sledované hodnoty. Schéma takejto matice a výsledný chord diagram vidíme na obr. 13.

Výhodou takto vytvorených dynamických chord diagramov je priama interakcia s reálnymi dátami a vstupnými entitami, ktoré sú v tabuľkovej podobe ukryté. Rovnako tak atraktívne kruhové usporiadanie je výhodou oproti iným bežne používaným grafickým vyjadreniam, na kto-

rých limity sme poukázali v časti „Cirkulárna vizualizácia“ nášho príspevku. Ako už bolo povedané, pri statickom vyjadrení chord diagramov treba čitateľom ukázať vhodný spôsob jeho interpretácie, resp. zacieliť jeho pozornosť na kľúčové dáta, ktoré mu chceme komunikovať (týka sa to predovšetkým veľkého množstva dát). Pri dynamickom vyjadrení je možné identifikovať aj dáta s najnižšími hodnotami. Aj z týchto dôvodov sa chord diagramy v súčasnosti úspešne etablovali ako plnohodnotný grafický prvok pri tvorbe podobne orientovaných štatistických dát.

		1	2	3	4	5		87	88	89	90	91	92
		Bratislava - Staré Mesto	Bratislava - Podunajské Biskupice	Bratislava - Ružinov	Bratislava - Vrakuňa	Bratislava - Nové Mesto	... ďalšie obce BA VÚC	Veľký Biel	Vlky	Zálesie	základné	stredné	vysoké
1	Bratislava - Staré Mesto	0	0	0	0	0	0	0	0	0	393	508	779
2	Bratislava - Podunajské Biskupice	0	0	0	0	0	0	0	0	0	375	375	376
3	Bratislava - Ružinov	0	0	0	0	0	0	0	0	0	710	884	1637
4	Bratislava - Vrakuňa	0	0	0	0	0	0	0	0	0	337	413	521
5	Bratislava - Nové Mesto	0	0	0	0	0	0	0	0	0	547	610	888
	... ďalšie obce BA VÚC	0	0	0	0	0	0	0	0	0	n	n	n
87	Veľký Biel	0	0	0	0	0	0	0	0	0	61	69	48
88	Vlky	0	0	0	0	0	0	0	0	0	34	25	15
89	Zálesie	0	0	0	0	0	0	0	0	0	102	46	44
90	základné	393	375	710	337	547	n	61	34	102	0	0	0
91	stredné	508	375	884	413	610	n	69	25	46	0	0	0
92	vysoké	779	376	1637	521	888	n	48	15	44	0	0	0



Obr. 13. Schéma matice a výsledný diagram pre diagramy „bez vzájomného vzťahu medzi entitami“

Interaktívnu verziu diagramu uvádzame na stránke sobdtn.sk (Bačík 2019f).

Záver

Cieľom predloženého príspevku bolo poukázať na možnosti zobrazenia dochádzkových dát pomocou cirkulárnej vizualizácie so zameraním na využitie tzv. chord diagramov. Impulzom pre napísanie nášho príspevku bola snaha o efektívne a atraktívne zobrazenie takýchto dát, s ktorými sa v geografickej, resp. demografickej praxi stretávame pomerne často. Efektívny spôsob vizualizácie umožní čitateľom získať špecifický pohľad na dáta a charakteristiky, ktoré sú v bežnej tabuľkovej podobe skryté. Inšpiráciu pre napísanie predloženého príspevku boli práce demografov (Abel, Sander 2014, Abel, Cohen 2019, Abel, Heo 2018), ktorí aplikovali túto formu grafického vyjadrenia na svetové migračné dáta. Práve takto orientované dáta, pri ktorých možno hovoriť o zdrojových a cieľových regiónoch, sú veľmi vhodným príkladom takéhoto využitia.

Cirkulárna vizualizácia sa stala pomerne dominantnou v ostatných rokoch. Samotné dôvody sú dobre zachytené v príspevku Lima (2017), ktorý na základe dostupných štúdií poukazuje na prirodzenú preferenciu ľudí vnímať okrúhle tvary ako pozitívny vnem zakorenený v ľudských organizmoch od narodenia. Popularita takéhoto grafického vyjadrenia je zrejmä aj s jej implementáciou vo viacerých štandardne používaných vizualizačných nástrojoch (napr. amCharts, Tableau, R, ZingChart, Highcharts, D3, Circos, atď.). V našom príspevku sme použili nástroj Circos na tvorbu neinteraktívnych chord diagramov, v prípade interaktívnych chord diagramov sme využili knižnicu D3, s ktorej využitím máme praktické skúsenosti aj z minulosti. Diagramy vytvorené pomocou týchto nástrojov možno rozdeliť do dvoch základných skupín – interaktívne (D3) a bez interaktivity (Circos). V oboch prípadoch však ide o atraktívne kruhové usporiadanie sledovaných entít (v našom prípade obce Bratislavského VÚC) a vzájomných vzťahov medzi nimi (dochádzkové dáta). Chord diagramy bez interaktivity sú vhodným nástrojom pri tvorbe prezentačných techník (postery, infografiky), pri ktorých chceme zaujať čitateľa už na prvý pohľad. Za istú nevýhodu môžeme považovať relatívnu komplikovanosť čitateľnosti týchto výstupov pre čitateľov, ktorí ešte nemajú s takýmito výstupmi bližšiu skúsenosť. Preto je vždy vhodné doložiť ako súčasť diagramu aj základný „návod“ ako postupovať pri jeho čítaní. Ďalšou nevýhodou môže byť znížená prehľadnosť čitateľnosti v prípade veľkého množstva vstupných dát. V takýchto prípadoch je však možné tieto diagramy upraviť tak, aby sme čitateľovu pozornosť zamerali na tie dáta, ktoré sú vzhľadom na výsledky nášho výskumu relevantné. Obe tieto nevýhody eliminuje tvorba interaktívnych chord diagramov. V týchto prípadoch je možné odčítanie aj tých najmenších hodnôt z diagramu, ktoré sú v statickej podobe viac-menej skryté. Samotná interaktivita je limitovaná znalosťami skriptovacích jazykov, ktorých jednotlivé techniky používajú tvorcovia výslednej prezentácie. Online chord diagramy pokrývajú spektrum od statických, až po dynamické animácie zachytávajúce dlhé časové obdobia (Abel, Sander 2014, Abel, Cohen 2019).

Dochádzkové dáta sú vzhľadom na svoj charakter veľmi vhodné na aplikovanie takejto vizuálnej formy. Tu sme znázornili jednak vzájomné dochádzkové väzby medzi sledovanými regiónmi, ako aj ďalšie dochádzkové atribúty, ktoré sú dostupné zo vstupných datasetov. V prípade migračných dát je možné aj zachytenie vnútroregionálnej migrácie (v takomto prípade by na diagonále vstupnej matice neboli nulové hodnoty, ale reálne hodnoty vnútroregionálnej migrácie). Vychádzajúc z konceptu popularity kruhového usporiadania entít a väzieb medzi nimi, je možné chord diagram použiť pri vizualizácii akýchkoľvek dát. Ako príklad možno uviesť aj aplikovanie chord diagramov na znázornenie výsledkov vzájomných zápasov medzi športovými tímami (športovcami)². Aj tu ide o obojsmernú interakciu a takto znázornené vzájomné súperenie môže byť obohatením športových štatistík.

Chord diagramy sú pomerne novým, avšak už pevne etablovaným grafickým prvkom pri vizualizácii dát. Ich využitie je veľmi rôznorodé a užívateľsky vysoko atraktívne, čo je predpokladom ich ďalšieho systematického využívania pri tvorbe rozličných printových a predovšetkým online interaktívnych aplikácií zameraných práve na vizualizáciu dát.

² Pre zaujímavosť sme spracovali vzájomné zápasy medzi všetkými hráčmi tenisového rebríčka ATP, ktorí boli aspoň 1 týždeň svetovou jednotkou. Výsledky sú dostupné na stránke: www.sodbtn.sk/atp/tenis.pdf (statická verzia), alebo www.sodbtn.sk/atp (interaktívna verzia)

Literatúra

- ABEL, G. J., COHEN, J. E. 2019: Bilateral international migration flow estimates for 200 countries. *Sci Data*, 6(82), 1-13. DOI: <https://doi.org/10.1038/s41597-019-0089-3>.
- ABEL G. J., HEO, N. 2018: Changing internal migration flows patterns in South Korea. *Regional Studies, Regional Science*, 5(1), 78-80. DOI: <https://doi.org/10.1080/21681376.2018.1431149>.
- ABEL, G. J., SANDER, N. 2014: Quantifying global international migration flows. *Science*, 343(6178), 1520-1522. DOI: <https://doi.org/10.1126/science.1248676>.
- ALI, S. M., GUPTA, N., NAYAK, G. K., LENKA, R. R. 2016: Big data visualization: Tools and challenges. In Niranjana, S. K., Manjunatha Aradhya, V. N. eds. *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*. India, Noida (IEEE), 2016, pp. 656-660. DOI: <https://doi.org/10.1109/IC3I.2016.7918044>.
- ANDERSEN, J. E., FUJI, H., MANABE, M., PENNER, R. C., SULKOWSKI, P. 2016: Enumeration of chord diagrams via topological recursion and quantum curve techniques. *Mathematical Physics. arXiv preprint arXiv:1612.05839*.
- BAČÍK, V. 2019a: Ako čítať chord diagram. *SOBDTN.sk*. Dostupné na: http://www.sodbtn.sk/dochadzka_ba/ako_citat_chord.png.
- BAČÍK, V. 2019b: Statické chord diagramy. *SOBDTN.sk*. Dostupné na: http://sodbtn.sk/dochadzka_ba/staticke_chord_diagramy.pdf.
- BAČÍK, V. 2019c: Vizualizácia dochádzky a odchádzky v podobe chord-diagramov. *SOBDTN.sk*. Dostupné na: http://www.sodbtn.sk/dochadzka_ba/.
- BAČÍK, V. 2019d: Odchádzajúci v mestských častiach BA medzi sebou. *SOBDTN.sk*. Dostupné na: http://sodbtn.sk/dochadzka_ba/chord18.php.
- BAČÍK, V. 2019e: Dochádzajúci v mestských častiach BA medzi sebou. *SOBDTN.sk*. Dostupné na: http://sodbtn.sk/dochadzka_ba/chord19.php.
- BAČÍK, V. 2019f: Odchádzka z obcí BA kraja do škôl. *SOBDTN.sk*. Dostupné na: http://www.sodbtn.sk/dochadzka_ba/chord6.php.
- BOSTOCK, M., OGIEVETSKY, V., HEER, J. 2011: D³ data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12), 2301-2309. DOI: <https://doi.org/10.1109/TVCG.2011.185>.
- BYLINSKII, Z., KIM, N. W., O'DONOVAN, P. et al. 2017: Learning visual importance for graphic designs and data visualizations. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 57-69. DOI: <https://doi.org/10.1145/3126594.3126653>.
- GRACE, S. 2018: Lessons from 'losses from the Russian campaign'. *The Art of Consequences*. Retrieved from: <https://edspace.american.edu/visualwar/minard/>.
- FLAJOLET, P., NOY, M. 2000: Analytic combinatorics of chord diagrams. In Krob, D., Mikhalev A. A., Mikhalev A. V. *Formal Power Series and Algebraic Combinatorics*. Berlin (Springer, Berlin, Heidelberg), pp. 191-201. DOI: https://doi.org/10.1007/978-3-662-04166-6_17.
- FRIENDLY, M. 2008: A brief history of data visualization. In Chen, Ch., Härdle, W., Unwin, A. eds. *Handbook of data visualization*. Berlin (Springer, Berlin, Heidelberg), pp. 15-56.
- HOLTEN, D. 2006: Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *IEEE Transactions on visualization and computer graphics*, 12(5), 741-748. DOI: <https://doi.org/10.1109/tvcg.2006.147>.
- KIRK, A. 2012: *Data Visualization: a successful design process*. Birmingham (Packt Publishing Ltd).
- KRZYWINSKI, M., SCHEIN, J., BIROL, I. et al. 2009: Circos: an information aesthetic for comparative genomics. *Genome research*, 19(9), 1639-1645. Retrieved from: <http://mkweb.bcgsc.ca/tableviewer/>. DOI: <https://doi.org/10.1101/gr.092759.109>.

- LIMA, M. 2017: *The book of circles: Visualizing spheres of knowledge*. New York (Princeton Architectural Press).
- PHAN, D., XIAO, L., YEH, R., HANRAHAN, P. 2005: Flow map layout. *IEEE Symposium on Information Visualization*, 219-224. DOI: <https://doi.org/10.1109/INFVIS.2005.1532150>.
- PURSEL, B. 2005: Information, People, and Technology. State College (The Pennsylvania State University). Retrieved from: <https://psu.pb.unizin.org/ist110/>.
- SANDER, N., ABEL, G. J., BAUER, R., SCHMIDT, J. 2014: Visualising migration flow data with circular plots. In Goujon, A., Streissnig, E. eds. *Working Papers*. Vienna (Vienna Institute of Demography).
- STOIMENOW, A. 2000: On the number of chord diagrams. *Discrete Mathematics*, 218(1-3), 209-233. DOI: [https://doi.org/10.1016/S0012-365X\(99\)00347-7](https://doi.org/10.1016/S0012-365X(99)00347-7).
- TUFTE, E. R. 2001: *The visual display of quantitative information* (Vol. 2). Cheshire, CT: Graphics press.
- WANG, C., YU, H., MA, K. L. 2008: Importance-driven time-varying data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6), 1547-1554. DOI: <https://doi.org/10.1109/TVCG.2008.140>.
- ZHOU, H., XU, P., YUAN, X., QU, H. 2013: Edge bundling in information visualization. *Tsinghua Science and Technology*, 18(2), 145-156. DOI: <https://doi.org/10.1109/TST.2013.6509098>.

Pod'akovanie: Príspevok vznikol vďaka projektu APVV-17-0079 „Analýza a prognóza demografického vývoja Slovenskej republiky v horizonte 2080: identifikácia a modelovanie dopadov na sociálno-ekonomickú sféru v rozličných priestorových mierkach“.

Adresy autorov

Mgr. Vladimír Bačík, PhD.

Katedra ekonomickej a sociálnej geografie, demografie a územného rozvoja

Prírodovedecká fakulta, Univerzita Komenského v Bratislave

Ilkovičova 6, Mlynská dolina,

842 15 Bratislava

Slovensko

vladimir.bacik@uniba.sk

Mgr. Dominika Chvosteková

Kristy Bendovej 4,

84102 Bratislava

Slovensko

dominika.chvostekova@gmail.com