

Seminář

Prostorová data a popis volební podpory

Mgr. Petr Voda, Ph.D.

Úvodem

- Prostorová analýza x analýza v prostoru
 - V **politologii** se prostorové hlasování (spatial voting) a prostorová analýza voleb (spatial analysis) týká myšleného politického prostoru
 - V **geografii** se prostorová analýza týká fyzického prostoru
- Kvantitativní
- V tradičním pojetí prostorové analýzy se neuplatňují jiné než kvantitativní metody

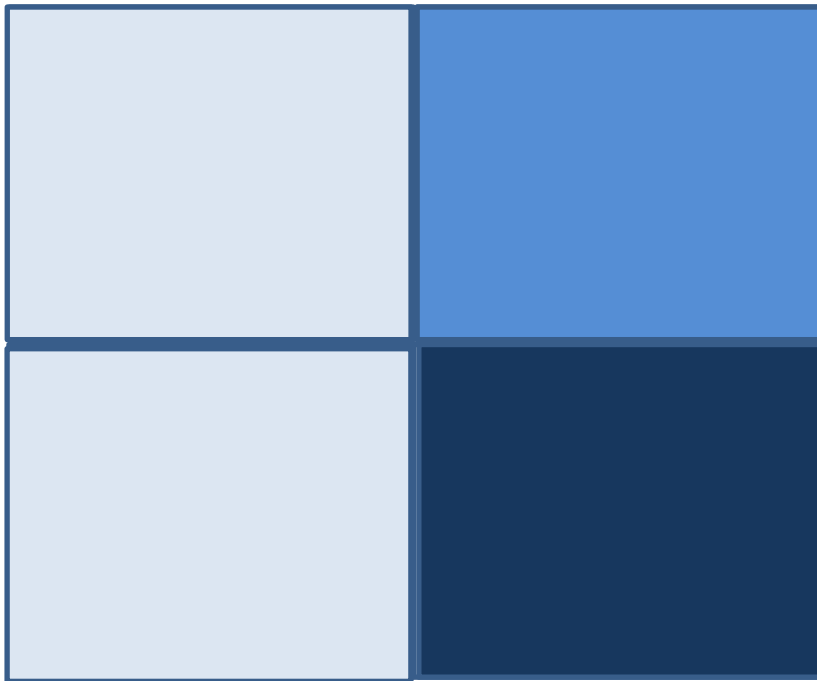
Popis x vysvětlení

- Popis
 - Statistické vlastnosti dat
 - Prostorové vlastnosti dat
- Vysvětlení
 - Proč je podpora strany rozložena popsáním způsobem?
 - **nutný teoretický předpoklad vztahů**

Povaha dat

- Agregovaná data
- Kardinální
- Prostorové
a časové zařazení

Princip vzniku agregovaných dat



- Různě vysokí lidé
- V několika místnostech
- Agregace dat
 - Jen jeden údaj za místnost
 - Různé situace mohou vést ke stejnému výsledku

Jméno	Adresa	věk	účast	strana	vyznání	EA
Karel J.	Údolní 5	18	Ano	ČSSD	ateista	zaměstnanec
Jana B.	Údolní 2	73	Ano	ČSSD	ateista	důchodce
Jiří K.	Údolní 12	34	Ne	-	ateista	OSVČ
Květa D.	Údolní 7	45	Ano	ANO	ateista	zaměstnanec
Tomáš V.	Marešova 3	98	Ano	ODS	ateista	důchodce
Marie H.	Marešova 7	26	Ne	-	katolík	zaměstnanec
Jan Z.	Marešova 5	22	Ano	TOP09	ateista	OSVČ
Eva A.	Jaselská 4	73	Ano	KSČM	ateista	důchodce
Zdeněk C.	Jaselská 9	55	Ne	-	ateista	zaměstnanec
Pavel N.	Jaselská 24	23	Ano	ANO	ateista	student
Anna R.	Jaselská 35	64	Ano	ANO	ateista	důchodce
Emil M.	Obilní trh 2	48	Ano	ČSSD	ateista	zaměstnanec
Lucie S.	Obilní trh 8	47	Ano	KDU	ateista	zaměstnanec
Milan T.	Obilní trh 13	49	Ne	-	katolík	OSVČ

Údolní

Marešova

Jaselská

Obilní trh

Jméno	Adresa	věk	účast	strana	vyznání	EA
Karel J.	Údolní 5	18	Ano	ČSSD	ateista	zaměstnanec
Jana B.	Údolní 2	73	Ano	ČSSD	ateista	důchodce
Jiří K.	Údolní 12	34	Ne	-	ateista	OSVČ
Květa D.	Údolní 7	45	Ano	ANO	ateista	zaměstnanec

$$(18+73+34+45)/4$$

$$3 \times \text{Ano} / 4$$

ulice	Průměrný věk	účast	ČSSD	ANO	KSČM	ODS	TOP	KDU	kato líci	zaměs tanci	OSV
Údolní	42,5	75	66,6	33,3	0	0	0	0	0	50	25
Marešova	48	66,6	0	0	0	50	50	0	33,3	33,3	33,3
Jaselská	50,1	75	0	66,6	33,3	0	0	0	0	25	0
Obilní trh	48	66,6	50	0	0	0	0	50	33,3	66,6	33,3

EKOLOGICKÁ CHYBA

- V procesu agregace se ztrácí informace
- Nelze hovořit o pozorovaných vztazích jako o platných pro individuální voliče
 - V našem případě: tam kde je nějaký katolík získala hlas KDU nebo TOP09. Přitom ale oba katolíci nevolili.
 - V reálném světě nejsme schopni věrohodně z agregovaných dat věrohodně rekonstruovat individuální vztahy

Specifika prostorových dat

- Autokorelace (více seminář 28. 4.)
 - „vše souvisí se vším, a co si je blíž, to spolu souvisí více“
 - Porušení předpokladu o nezávislosti pozorování
- Nestacionarita (více seminář 5. 5.)
 - Volební chování jedné společenské skupiny se může v prostoru lišit (katolíci ve Valašských Klouboucích x katolíci v severních Čechách, podnikatelé v Praze x podnikatelé na Svitavsku)

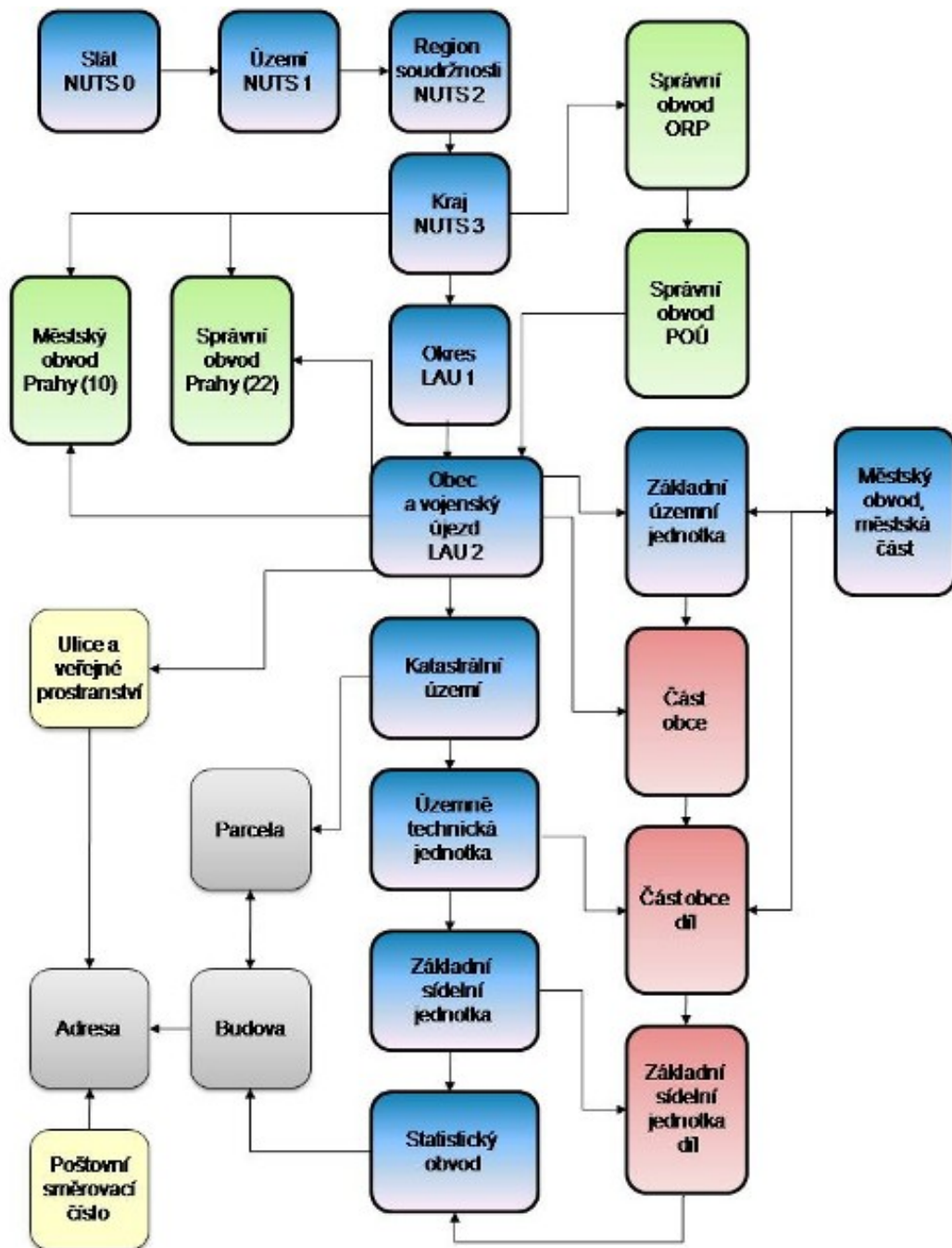
Vlastnosti dle reprezentace prostoru

- (izolované) objekty x (spojitá) pole
- Objekty: např. obce
- Pole: např. vzduch

- Reprezentace
 - Body
 - Linie
 - Polygony

Velikost polygonu

- Malý region = vysoká homogenita/
vysoký „šum“
 - Velký region = nízká homogenita/
nízký šum
- Funkční x administrativní region



- <http://apl.czso.cz/irso4/cisel.jsp>

- Viz

- <http://www.cuzk.cz/Uvod/Produkty-a-sluzby/RUIAN/2-Poskytovani-udaju-RUIAN-ISUI-VDP/Ciselniky-ISUI/Nizsi-uzemni-prvky-a-uzemne-evidencni-jednotky.aspx>



Zdroj: [google.maps.com](https://www.google.com/maps)

Rozdíly mezi měřítky

	nezaměstnanost		vš		Počet obyvatel		OSVČ		
	min	max	min	max	min	max	min	max	
Brno	9,4		20,6		385913		14,19		1
Městské části	5,8	11,9	11,1	32,5	577	64316	9,2	18,3	29
Části	0	23	0	39	5	26781	6	75	59
ZSJ	0	100	0	75	0	12836	0	100	264

- kvůli značným rozdílům ve velikosti jednotek je obvykle vhodné používat váhy

Důsledky „měření“ (sběru dat)

- V socio-ekonomické analýze obvykle není problém s chybějícími daty za místa
- Problém s chybějícími daty pro čas (mnoho údajů je zjišťováno jen z cenzu)
- Bojkot sčítání (např. Řekové v Albánii, Albánci v Makedonii a Srbsku, ...)
 - V ČR otázka víry a vyznání v roce 2011

Shrnutí

- Agregovaná data
 - Nebezpečí ekologické chyby
- Kardinální proměnné
 - Možnosti pro využití řady statistických nástrojů
- Prostorová data
 - Narušení obvyklých předpokladů
 - Otázka měřítkové úrovně
 - Otázka spolehlivosti dat

Data pro analýzu

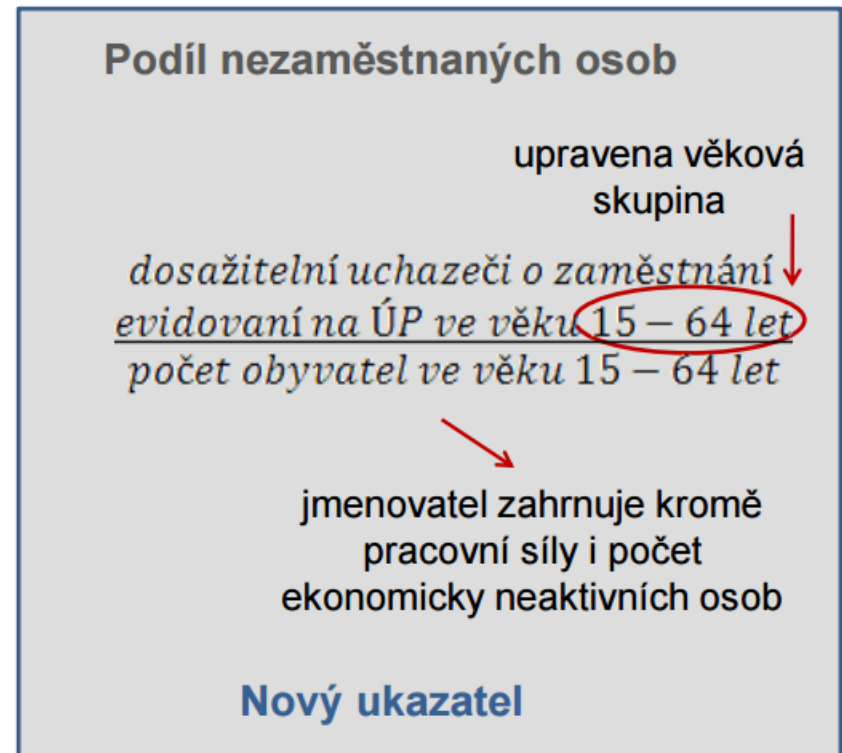
- Volební výsledky (předmět výzkumu, závisle proměnná)
 - Nejlépe jako % platných hlasů pro stranu v území
 - V longitudálním výzkumu % z oprávněných voličů – odpadá problém s účastí
- socio-ekonomická data (nezávisle proměnné)
 - Možnost vypočítat celou řadu indikátorů
 - Různá pravidla pro různé ukazatele

Ukazatele

- Ekonomická aktivita:
 - Míra nezaměstnanosti
 - Podíl pracujících v sektorech ekonomiky (zemědělství, průmysl, služby)
 - Podíl pracujících v soukromém x veřejném sektoru
 - Podíl osob s různým postavením v zaměstnání (zaměstnanec, zaměstnavatel, osvč)
- Vzdělání
 - Podíl osob s VŠ vzděláním
- Víra
 - Podíl osob s určitým vyznáním

Míra nezaměstnanosti

- Z MPSV
- Obecná míra nezaměstnanosti



Míra nezaměstnanosti

- Z ČSÚ do roku 2012
- Míra registrované nezaměstnanosti –
 - v čitateli počet neumístěných uchazečů o zaměstnání a ve jmenovateli disponibilní pracovní síla, tj. zaměstnaní z VŠPS + neumístění uchazeči
 - Od července 2004:
 - v čitateli počet dosažitelných neumístěných uchazečů o zaměstnání a ve jmenovateli součet počtu zaměstnaných z VŠPS, pracujících cizinců registrovaných na úřadech práce, nebo s platným povolením k zaměstnávání, či živnostenským oprávněním a počtu dosažitelných neumístěných uchazečů o zaměstnání

Míra nezaměstnanosti

- Ze sčítání lidu
- = $\text{Nezaměstnaní} / \text{ekonomicky aktivní} * 100$

Vzdělání

- Typ nejvyššího ukončeného vzdělání / obyvatelstvo starší 15 let *100
- Osoby s vyšším odborným vzděláním se obvykle počítají jako osoby s maturitou (ne VŠ)

Víra

- Kategorie Věřící ve sčítání lidu 2011 obsahuje všechny (včetně rytířů Jedi a vyznavačů špagetového monstra)
- Vhodné pracovat jen s jasnými kategoriemi
 - Počet osob s římskokatolickým vyznáním/počet obyvatel celkem *100

Další ukazatele

- Podíl rodáků (narození v obci obvyklého bydliště/obyvatelé * 100)
- Podíl osob nad 65 let
- Podíl osob žijících v bytových domech
- ...
- <http://apl.czso.cz/iSMS/ukazvyb.jsp>
- <https://vdb.czso.cz/vdbvo2/faces/cs/index.jsf?page=metodika-uvod>

Popis rozložení volební podpory

- Průměr + směrodatná odchylka
- Variační koeficient
- Giniho koeficient
- Území (stabilní) volební podpory a indikátory s ním spojené

Úvod

- Odpověď na otázku JAK?
- Kolik získaly strany hlasů?
- Jaké byly rozdíly mezi zisky stran?
- Jak je koncentrovaná podpora stran ve volbách?
- ...

Proč?

- Účelem deskriptivní statistiky je zjistit vlastnosti proměnné
 - (Porozumět používaným datům)
- Vlastnosti proměnné mají důsledek pro další analýzy
- A pro interpretaci výsledků analýz
- (vytváření agregovaných dat)

Průměr + sm.odch.

- Průměr - Hledáme hodnotu, která nejlépe reprezentuje proměnnou
- Samotná střední hodnota poskytuje značně redukovanou (a zkreslenou) představu o vlastnostech proměnné
- Míra variability doplňuje informaci, jak dobře střední hodnota reprezentuje všechny případy
- Ukazuje, jak moc se mezi sebou liší hodnoty proměnné

Variační koeficient

- =směrodatná odchylka/průměr
- V prostorové analýze využíván jako míra koncentrace
- Neznamená ale míru územní koncentrace!!!
- Nebere v potaz rozložení hodnot v prostoru
- 0 – velmi malé rozdíly v hodnotách proměnné
- Nemá pevnou horní hranici

Giniho koeficient

- Rozdíl skutečného a ideálně rovnoměrného rozložení
- Není obsažen v spss ani excelu
- Seřadit jednotky dle relativního zisku
- Z absolutních hodnot vypočítat kumulovaná procenta
- Vložit do wessa.net
(<http://www.wessa.net/co.wasp>)

Území volební podpory

- Přehlednost, zohlednění populační velikosti
- Nevhodné pro lokální strany (např. SMK v SVK)
- Doplnující indikátory
 - Míra úspěšnosti: kolikrát je podpora strany vyšší v (jádru) území volební podpory oproti zbytku území

Popis vývoje volební podpory

- Bazické a řetězové indexy
 - Index volební úspěšnosti
 - Index volební stability
- Korelační koeficient
- Území stabilní volební podpory

Bazické a řetězové indexy

- Řetězový index
 - Volby 2013/volby 2010; volby2010/volby2006 ...
 - Ukazuje postupný vývoj
- Bazický index
 - Volby 2013/volby2002, volby2010/volby 2002
 - Ukazuje změnu stav oproti stanovenému základu

Korelační analýza

- Míra souvislosti mezi dvěma proměnnými
- Lineární vztah
- Možnost vytvořit (libovolně) velkou matici
 - Souvislost následujících voleb i souvislost vzdálených voleb

Území stabilní volební podpory

- Strany s koncentrovanou stabilní podporou X strany s nekoncentrovanou stabilní podporou
- Výpočet procenta hlasů přítomných v ÚSVP v každých volbách

Analýza normálního rozdělení

- Průměr, medián
- Směrodatná odchylka
- Kvartily
- Šikmost, strmost

- testy

Proč?

- Můžeme použít proměnné v další analýze?