

**Univerzita Palackého v Olomouci**  
**Přírodovědecká fakulta**  
**Katedra geoinformatiky**

**PROSTOROVÉ HODNOCENÍ DAT  
VOLEBNÍCH VÝSLEDKŮ**

**Diplomová práce**

**Přemysl BONK**

**Vedoucí práce Mgr. Karel Macků Ph.D.**

**Olomouc 2024**  
**Geoinformatika a kartografie**

## **ANOTACE**

Diplomová práce se zabývá prostorovým hodnocením volebních výsledků pomocí prostorových statistických metod. Teoretická část diplomové práce má poskytnout základní vhled do problematiky a stručnou charakteristiku metod hodnotící volební výsledky ze statistického a prostorového hlediska. Samotné hodnocení bylo provedeno v praktické části práce na základě sestavené datové sady obsahující samotná data volebních výsledků a vybrané sociální, ekonomické a demografické charakteristiky.

V rámci diplomové práce bylo zpracováno pět analýz pro volby do poslanecké sněmovny 2021 a voleb prezidenta republiky 2023 České republiky. Zpracované analýzy se zabývaly hodnocením vícerozměrné odlehlosti vstupních atributů, využitím voličských průkazů, popisem variabilního chování volebních výsledků využitím GWR modelů, predikcí na základě modelů logistické regrese a shlukovou analýzou. Veškeré analýzy byly provedeny nad administrativními celky základních územních jednotek.

Hlavním výstupem provedených analýz volebních výsledků je deset mapových kompozic, jejichž součástí jsou kromě map také tabulky, grafy apod., poskytující další informace o výsledcích provedených analýz. Ty poskytují komplexní přehled volebního chování zúčastněných voličů a jejich preferencí v kontextu proběhlých voleb.

## **KLÍČOVÁ SLOVA**

vícerozměrná odlehlost; voličské průkazy; GWR; logistická regrese; shluková analýza

Počet stran práce: 72

Počet příloh: 20 (z toho 17 elektronických, 1 volná a 2 vázané)

## **ANOTATION**

The diploma thesis deals with the spatial evaluation of election results data using spatial statistical methods. The theoretical part of the thesis provides a basic insight into the issue and a brief characterization of methods evaluating election results data from a statistical and spatial point of view. The practical part of the thesis aims to evaluate the election result data based on a dataset containing the election results data itself and selected social, economic and demographic characteristics.

Five analyses were developed as part of the thesis for the 2021 Chamber of Deputies and 2023 Presidential elections in the Czech Republic. The analyses developed dealt with the evaluation of multivariate outliers of input attributes, the use of voter ID cards, the description of the variable behaviour of election results using GWR models, prediction based on logistic regression models and cluster analysis. All analyses were conducted over administrative units of basic territorial units.

The main output of the analyses of electoral results are ten map outputs, which, in addition to maps, also include tables, graphs, etc., providing additional information on the results of the analyses. The results of each analysis provide a comprehensive overview of voting behaviour and preferences in the context of the chosen elections.

## **KEYWORDS**

multivariate outlier; voter cards; GWR; logistic regression; cluster analysis

Number of pages 72

Number of appendixes 20

**Prohlašuji, že**

- bakalářskou/diplomovou práci včetně příloh, jsem vypracoval(a) samostatně a uvedl(a) jsem všechny použité podklady a literaturu,

- jsem si vědom(a), že na moji bakalářskou/diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevydělečně, ke své vnitřní potřebě, bakalářskou/diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, že údaje o mé bakalářské/diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé bakalářské/diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé bakalářské/diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne 6. 5. 2024

Bc. Přemysl BONK  
podpis autora

Děkuji vedoucímu práce Mgr. Karlu Macků, Ph.D., za odborné vedení, cenné rady a užitečné připomínky při konzultacích k vypracování diplomové práce. Dále bych rád poděkoval, Mgr. Jakubu Lyskovi, Ph.D., za odbornou pomoc při zpracování analýz volebních výsledků. Za konzultace a doporučení při zpracování mapových výstupů děkuji, Mgr. Radku Barvířovi, Ph.D. Nakonec bych chtěl poděkovat své rodině a blízkým za nezjištnou podporu v době celého studia.

# UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Akademický rok: 2022/2023

## ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Bc. Přemysl BONK  
Osobní číslo: R220005  
Studijní program: N0532A330009 Geoinformatika a kartografie  
Téma práce: Prostorové hodnocení dat volebních výsledků  
Zadávající katedra: Katedra geoinformatiky

### Zásady pro vypracování

Cílem diplomové práce je zpracovat analýzu dat volebních výsledků s využitím vhodných prostorově-statistických metod. Nejprve student sestaví dataset vhodných, volně dostupných dat – primárně se zaměří na data voleb do Poslanecké sněmovny Parlamentu České republiky a na prezidentské volby. Volební data budou dále rozšířena o vhodné socioekonomické a demografické ukazatele, které následně poslouží pro zkoumání závislostí mezi volebními výsledky a charakteristikami území. Následně student aplikuje vhodné prostorově-statistické metody za účelem hledání odlehklých hodnot, identifikaci shluků podobných hodnot, popsání vnitřních závislostí v datech a sestavení modelů popisujících vztahy mezi volebními výsledky a socio-ekonomicko-demografickou situací v místě volby. Důraz bude kladen především na využití prostorově založených statistických metod. Závěrem student vyhodnotí vhodnost zvolených metod, která mohou být doporučením pro budoucí práce s tímto typem dat a jednotlivé výsledky vizualizuje. Výsledná zjištění budou hodnocena a interpretována v součinnosti s politologem.

Shromážděná data a všechny výstupy student přiloží k práci v digitální formě. Údaje o všech datových sadách, které vytvořil nebo získal v rámci práce, do Metainformačního systému katedry geoinformatiky a současně provede zálohu údajů ve formě validovaného XML souboru. O práci student vytvoří webové stránky, které budou v souladu s pravidly dostupnými na stránkách katedry a poster ve formátu A2. Na závěr práce připojí student jednostránkové resumé v anglickém jazyce. Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data, XML soubor) bude odevzdána v digitální podobě na paměťovém médiu a text práce bude spolu s vybranými přílohami odevzdán ve dvou svazcích na sekretariát katedry.

Rozsah pracovní zprávy: max. 50 stran  
Rozsah grafických prací: dle potřeby  
Forma zpracování diplomové práce: elektronická

### Seznam doporučené literatury:

- Lysek, J., & Macků, K. (2022). Continuity in Discontinuity? Spatial Patterns of Electoral Behaviour in the 2021 Czech Parliamentary Elections. *Czech Journal of Political Science*, 2022(3). <https://doi.org/10.5817/PC2022-3-260>
- Lysek, J., Pánek, J., & Lebeda, T. (2021). Who are the voters and where are they? Using spatial statistics to analyse voting patterns in the parliamentary elections of the Czech Republic. *Journal of Maps*, 17(1), 33–38. <https://doi.org/10.1080/17445647.2020.1819901>
- Maškarinec, P. (2017). A Spatial Analysis of Czech Parliamentary Elections, 2006–2013. *Europe-Asia Studies*, 69(3), 426–457. <https://doi.org/10.1080/09668136.2017.1313962>
- Mészáros, J., Solymosi, N., & Speiser, F. (2007). Spatial distribution of political parties in Hungary 1990–2006. *Political Geography*, 26(7), 804–823. <https://doi.org/10.1016/j.polgeo.2007.06.002>
- Zarycki, T. (2015). The electoral geography of Poland: between stable spatial structures and their changing interpretations. *Erdkunde*, 69(2), 107–124. <https://doi.org/10.3112/erdkunde.2015.02.02>

Vedoucí diplomové práce: **Mgr. Karel Macků, Ph.D.**  
Katedra geoinformatiky

Datum zadání diplomové práce: **9. prosince 2022**  
Termín odevzdání diplomové práce: **9. května 2024**

# ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Ing. Petr Bábek**  
Ročník: **1922022**  
Studijní program: **1922A330608 Geoinformatika a kartografie**  
Téma práce: **Porovnání bodových dat vektorních výkresů**  
Katedra: **Geoinformatika**

## Zásady pro vypracování

Čtenářům této práce je doporučeno seznámit se s obsahem předložených zdrojů informací a zejména s metodami jejich zpracování a prezentace. Účelem této práce je porovnání bodových dat vektorních výkresů a jejich prezentace v prostředí GIS. Práce je určena pro studenty, kteří se zabývají geoinformatikou a kartografií. Práce je rozdělena do dvou částí: teoretické a praktické. Teoretická část obsahuje popis principů vektorních výkresů a jejich prezentace v prostředí GIS. Praktická část obsahuje popis metod porovnání bodových dat vektorních výkresů a jejich prezentace v prostředí GIS. Práce je určena pro studenty, kteří se zabývají geoinformatikou a kartografií. Práce je rozdělena do dvou částí: teoretické a praktické. Teoretická část obsahuje popis principů vektorních výkresů a jejich prezentace v prostředí GIS. Praktická část obsahuje popis metod porovnání bodových dat vektorních výkresů a jejich prezentace v prostředí GIS.

Pracovní postup práce: **elektronická**  
Forma předání práce: **elektronická**

LS.

**doc. RNDr. Martin Kubala, Ph.D.**  
děkan

**prof. RNDr. Vilém Pechanec, Ph.D.**  
vedoucí katedry

V Olomouci dne 1. září 2023

# OBSAH

<b>SEZNAM POUŽITÝCH ZKRATEK .....</b>	<b>9</b>
<b>ÚVOD .....</b>	<b>10</b>
<b>1 CÍLE PRÁCE.....</b>	<b>11</b>
<b>2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY .....</b>	<b>12</b>
2.1 Volební geografie .....	14
2.2 Prostorové analýzy volebních výsledků.....	15
2.3 Odborné práce zabývající se podobným tématem .....	17
<b>3 METODY A POSTUP ZPRACOVÁNÍ .....</b>	<b>19</b>
3.1 Použité metody .....	19
3.1.1 Mahalanobisova vzdálenost.....	19
3.1.2 Lineární regrese a GWR.....	19
3.1.3 Logistická regrese.....	20
3.1.4 Shluková analýza .....	20
3.1.5 Metody výpočtu volebních ukazatelů.....	21
3.2 Použitá data .....	21
3.2.1 Data volebních výsledků.....	21
3.2.2 Data SED ukazatelů.....	21
3.3 Použité programy.....	22
3.3.1 ArcGIS Pro.....	22
3.3.2 QGIS.....	22
3.3.3 MS Excel .....	22
3.3.4 RStudio .....	22
3.3.5 Inkscape.....	22
3.4 Postup zpracování .....	23
<b>4 VLASTNÍ ŘEŠENÍ.....</b>	<b>24</b>
4.1 Sestavení datové sady .....	24
4.1.1 Data volebních výsledků.....	24
4.1.2 Data SED ukazatelů.....	26
4.1.3 Exploratorní analýza .....	28
4.2 Aplikace prostorových analýz .....	36
4.2.1 Mahalanobisova vzdálenost.....	36
4.2.2 Analýza využití voličských průkazů .....	44
4.2.3 Geograficky vážená regrese .....	46
4.2.4 Predikce s využitím logistické regrese.....	53
4.2.5 Shluková analýza .....	56
<b>5 VÝSLEDKY .....</b>	<b>63</b>
5.1 Sestavení datové sady .....	63
5.2 Aplikace prostorových analýz .....	63
<b>6 DISKUSE.....</b>	<b>68</b>
<b>7 ZÁVĚR.....</b>	<b>71</b>
<b>POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE</b>	
<b>PŘÍLOHY</b>	



## SEZNAM POUŽITÝCH ZKRATEK

<b>Zkratka</b>	<b>Význam</b>
CSV	Comma Separated Values
ČSSD	Česká strana sociálně demokratická
ČSÚ	Český statistický úřad
GPKG	OGC GeoPackage
GWR	Geographically weighted regression
HHI	Herfindahl-Hirschmanův Index
KDU–ČSL	Křesťanská demokratická unie – Česká strana lidová
K-S test	Kolmogorovův–Smirnovův test
KSČM	Komunistická strana Čech a Moravy
LISA	Local Indicators of Spatial Association
MPSV	Ministerstvo práce a sociálních věcí
MU	Masarykova Univerzita
ODS	Občanská demokratická strana
SED	socio-ekonomicko-demografické
SHP	Esri Shapefile
SLBD	Sčítání lidí domů a bytů
SPD	Svoboda a přímá demokracie
UPOL	Univerzita Palackého v Olomouci
XLSX	Microsoft Excel Open XML Spreadsheet

## ÚVOD

V demokratickém státním zřízení, kde mezi sebou různé politické subjekty vedou boj o nejvyšší volební zisk, je důležité znát postoje, preference a přání obyvatel s volebním právem. Z tohoto důvodu je v současné době stále častěji skloňovaným tématem analýza volebních výsledků, mimo jiné i z prostorového hlediska, přestože to nemusí být na první pohled zřejmé. Prostorovým umístěním administrativní jednotky je ovlivněna celá řada sociálních ekonomických a demografických faktorů, jenž mají vliv na volební zisk nebo ztrátu v konaných volbách. Nesprávná interpretace analýz volebních výsledků pak vede ke špatným rozhodnutím a ztrátě stávajících voličů. Z tohoto důvodu je stále více skloňovaným tématem politický marketing, mezi jehož postupy se řadí také prostorové analýzy volebních výsledků. Příkladem mohou být současné výjezdy ministra vnitra ČR Víta Rakušana z hnutí Starostové a nezávislí do lokalit, kde hnutí zaznamenalo ve volbách do Poslanecké sněmovny Parlamentu ČR (dále jen sněmovních volbách) nejhorší volební výsledky (Poláková-Uvírová 2024).

Hlavní motivací této práce je aplikace pokročilých i základních prostorových statistických metod pro hodnocení volebních výsledků voleb do Poslanecké sněmovny 2021 a voleb prezidenta republiky 2023. S využitím zmíněných metod je možné popsat prostorové vzory a trendy, které stojí za úspěchy nebo porážkami politických subjektů v regionech České republiky. Výsledky a výstupy zpracovaných analýz mohou přinést odpovědi na dlouho kladené otázky z řad politologů nebo posloužit jako příručka pro jejich využití v budoucích volbách.

# 1 CÍLE PRÁCE

Cílem diplomové práce je zpracování analýzy dat volebních výsledků s využitím vhodných prostorových statistických metod.

V první části práce bude sestavena datová sada volně dostupných dat. Primárním cílem je shromáždění dat výsledků voleb do Poslanecké sněmovny Parlamentu ČR a voleb prezidenta republiky. Důraz bude kladen především na shromáždění dat volebních výsledků za roky 2021 a 2023. Dále bude datová sada doplněna o vhodné socio-ekonomicko-demografické (dále jen SED) ukazatele. SED ukazatele budou využity ke zkoumání závislosti mezi volebními výsledky a charakteristikami území. Zdrojem těchto ukazatelů bude především Sčítání lidu, domů a bytů (dále jen SLBD 2021).

V druhé části diplomové práce budou aplikovány vhodné prostorové statistické metody, jejichž účelem bude hledání odlehklých hodnot nebo identifikace shluků podobných hodnot nebo sestavení modelů, které budou popisovat vztahy mezi volebním výsledkem a SED situací. Hlavní důraz bude při zpracování kladen především na využití prostorových statistických metod. Ke zpracování analýz budou využity geografické informační systémy a balíky funkcí vytvořené v jazyce R.

Ve třetí části diplomové práce budou výsledky analýz vhodně vizualizovány. Dále bude provedeno hodnocení a interpretace zjištěných poznatků v součinnosti s politologem. V této části budou využité metody zhodnoceny z hlediska vhodnosti využití pro analýzu volebních výsledků.

Výsledky a postupy využité při zpracování diplomové práce mohou být inspirací nejen pro politologické, ale i další výzkumy, jejichž součástí je geografický aspekt.

## 2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Volby a volební výsledky jsou obecně chápány jako pojmy, kterými se zabývá zejména politologie, popřípadě politická geografie. S rozvojem moderních technologií a výpočetních metod byly zpřístupněny nové možnosti analýzy volebních výsledků. Geoinformační technologie a metody prostorové statistiky jsou prostředkem k vysvětlení prostorových vztahů, které za předpokladu jejich správné interpretace, umožňují hlubší pochopení vnitřních souvislostí vedoucích k daným volebním výsledkům a volebnímu chování.

### **Shrnutí politického vývoje na území ČR od roku 1989**

Pádem komunistického režimu v listopadu roku 1989 byl politický vývoj tehdejšího Československa, stejně jako v ostatních bývalých komunistických státech východní Evropy, charakteristický repluralizací politického prostředí a ukončením vedoucího postavení komunistické strany (Wolchik 1993). V roce 1993 došlo k rozdělení tehdejší federativní republiky na dva nové státy v podobě České a Slovenské republiky, jež bylo iniciováno Václavem Klausem a Vladimírem Mečiarom (Šimíček 2014), čímž započal samostatný vývoj českého politického prostředí. Jednou z nejkontroverznějších událostí od té doby bylo podepsání tzv. opoziční smlouvy roku 1998, umožňující menšinovou vládu ČSSD bez výrazného odporu ze strany ODS. Jedním z důvodů kritiky uzavřené opoziční smlouvy je odsunutí tehdejších menších politických subjektů, např. KDU–ČSL nebo Unie svobody, na vedlejší kolej (Švec 2018). Po zrušení opoziční smlouvy oběma stranami, které ji dříve podepsaly, je vývoj českého politického prostředí v letech 2002–2014 charakteristický střídáním se vlád ODS nebo ČSSD s koaličními partnery. Do tohoto období spadají také dvě úřednické vlády, v jejichž čele stáli Jan Fisher a Jiří Rusnok (Vláda ČR 2024). Vítězem sněmovních voleb v roce 2013 byla ČSSD s 20,45 % hlasů, která utvořila pozdější vládní koalici s hnutím ANO 2011 s 18,65 % hlasů (Bahounková 2013), což byl počátek působení tohoto hnutí na české politické scéně. V následujících volbách do Poslanecké sněmovny ČR roku 2017 hnutí ANO 2011 zvítězilo se ziskem 29,64 % hlasů, a opět vytvořilo koalici s ČSSD (Hájek a Švec 2017).

V kontextu prezidentských voleb je důležitým milníkem zavedení přímé volby hlavy státu od roku 2013 (Vláda ČR 2010). Do tohoto roku byly oba předchozí prezidenti Václav Havel a Václav Klaus zvoleni nepřímou volbou prostřednictvím parlamentu ČR Vítězem v prvních prezidentských volbách se ziskem 54,80 % hlasů se stal občanský kandidát Miloš Zeman se ziskem 54,8 % hlasů, který ve druhém kole porazil Karla Schwarzenberga, a nahradil tak posledního prezidenta zvoleného nepřímou volbou, Václava Klause (Charvát 2014). V historicky druhých přímých volbách hlavy státu opět zvítězil Miloš Zeman s 51,36 % hlasů, tentokrát nad Jiřím Drahošem, jenž v minulosti zastával funkci předsedy Akademie věd ČR (Dolejší 2018).

Po volbách do poslanecké sněmovny v roce 2021 byla nastolena nová politická situace. První výraznou změnou byl neúspěch stran ČSSD a KSČM, které nepřekročily potřebnou hranici 5 %. Tento neúspěch byl způsoben spoluúčastí na vládě s hnutím ANO 2011 v předchozích dvou volebních obdobích a přechodem části bývalých voličů k hnutí ANO 2011, které začalo prosazovat více levicové hodnoty (Lysek a kol. 2022). Zajímavým poznatkem je situace okolo preferenčních hlasů, jež byla způsobena úspěchem hnutí Starostové a nezávislí a strany KDU–ČSL v zisku mandátů, a naopak velmi slabý výsledek u Pirátské strany. Dalším zjištěním je změna voličských preferencí, která byla způsobena pandemií COVID–19 (STEM 2021).

Taktéž v průběhu prezidentských voleb roku 2023 bylo zaznamenáno několik důležitých změn. Poprvé v historii přímé volby prezidenta odstoupil kandidát na prezidenta Josef Středula, jenž splnil požadavky k účasti v prezidentských volbách. Další zajímavostí

je dosud nejvyšší volební účast v historii přímé volby prezidenta republiky ČR, když právo volit využilo 68,24 % oprávněných voličů (ČSÚ 2023).

### **Charakteristika významných politických stran a hnutí v České republice**

V kontextu charakteristiky současných politických stran nebo hnutí a cílových skupin voličů je příhodné využití klasického členění politického spektra na pravici střed a levice. K tomuto dělení je dále často přidávána konzervativní či liberální orientace daného subjektu. Mezi pravicové politické subjekty jsou obvykle řazeny strany, jejichž prioritou je především důraz na jednotlivce, svoboda podnikání, nižší daně a obecně slabší význam státních institucí. Klasickými pravicovými voliči jsou např. podnikatelé a obyvatelstvo s nadprůměrnými příjmy. Charakteristikou uskupení hlásících se k levicové části politického spektra jsou požadavky na vyšší daně, sociální jistoty a silnou pozici státu. Levicové politické subjekty se zaměřují především na podporu nižší střední třídy, zaměstnanců státu, nebo také seniorů. Středovými subjekty politického spektra je prosazována umírněná politika kombinující oba vymezené směry (Strmiska 2017).

Klasickým příkladem pravicové strany je ODS (Interview Plus 2019), která v roce 2021 kandidovala v rámci koalice SPOLU se stranou TOP 09 a KDU–ČSL, přičemž obě tyto strany jsou řazeny do středopravého spektra. Nejstarší levicovou stranou české republiky je ČSSD a v krajně levicové části spektra také KSČM. Těmito stranami však nebyly splněny požadavky pro vstup do poslanecké sněmovny Parlamentu ČR. Koalice Pirátů a Starostů, která zaznamenala ve volbách 2021 třetí nejvyšší volební zisk, se hlásí ke středovému programu (Piráti a Starostové 2023). V deklarovaném programu je uvedeno odpovědné hospodaření s veřejnými financemi nebo dostupnost a efektivita státních služeb. Koalice Pirátů a Starostů chce tímto programem oslovit především liberální a progresivní voliče (Doubřavová a Dohnalová 2021)

V posledních letech je politická situace v České republice charakteristická nárůstem úspěšnosti nových politických subjektů s populistickými tendencemi (Lysek a kol 2022). Příkladem těchto subjektů jsou hnutí ANO 2011 a SPD, jejichž charakteristikou je líbivá, ale ne vždy odpovědná politika, například v kontextu rostoucích státních výdajů při absenci odpovídajících finančních zdrojů. Jedním z následků nárůstu populismu je přechod voličů od levicových politických stran ČSSD a KSČM k hnutím ANO 2011 a SPD.

### **Volební systém České republiky**

Od vzniku samostatné České republiky jsou v pravidelných intervalech opakovány volby do následujících institucí (Ministerstvo vnitra České republiky 2024):

- volby do Poslanecké sněmovny Parlamentu ČR (každé 4 roky),
- volby do Senátu Parlamentu ČR (každé 2 roky pro 1/3 senátu),
- volby do Evropského parlamentu (od roku 2004, každých 5 let),
- volby do zastupitelstev krajů (každé 4 roky),
- volby do zastupitelstev obcí (každé 4 roky),
- volba prezidenta ČR (od roku 2013, každých 5 let).

Ačkoliv je v České republice organizována celá řada voleb, dlouhodobě nejpobulárnějšími jsou mezi lidmi volby do Poslanecké sněmovny Parlamentu ČR a volba prezidenta ČR (Novinky.cz 2023). Z tohoto důvodu bylo rozhodnuto o zpracování analýz volebních výsledků především pro tyto volby za roky 2021 a 2023, neboť se jedná o nejpobulárnější a zároveň poslední proběhlé volby k datu vypracování diplomové práce.

## 2.1 Volební geografie

Volební geografii lze podle Gregoryho a kol. (2011) chápat jako „...vědní disciplínu, která studuje geografické aspekty průběhů a výsledků voleb. Hlavním objektem studia jsou pak faktory volebních výsledků ve vymezených územních celcích.“

### Faktory ovlivňující výsledky voleb

Volby do poslanecké sněmovny i volby prezidenta republiky jsou ovlivněny řadou faktorů. Stranická příslušnost je podle Bartelse (2000) důležitým faktorem, neboť představuje jádro voličů strany. V současné době je ale významnost tohoto faktoru na ústupu.

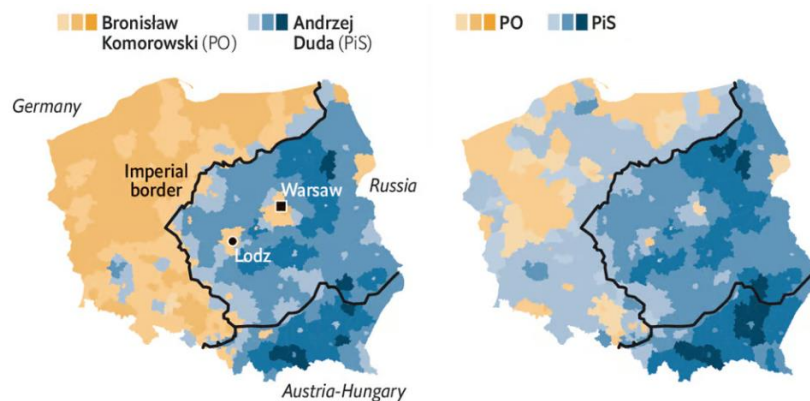
V kontextu analýzy volebních výsledků jsou důležité socioekonomické faktory, kterými je chování voličů ovlivněno. Podle Wiatra (1962) existuje korelace mezi volebním chováním a socioekonomickými faktory, nicméně volební chování mohou dále ovlivnit např. politické přesvědčení, psychický stav apod. Vliv různých socioekonomických faktorů je shrnut v publikaci *Factors Influencing Voting Decision: A Comprehensive Literature Review* (Kulachai 2023). Příkladem socioekonomických faktorů, které ovlivňují volební chování, jsou příjem, vzdělání, věk, zkušenost se zdravotní péčí nebo sociální a kulturní identita.

Volby jsou ovlivněny také fyzicko-demografickými faktory, jejichž působení na volební chování je popsáno v článku od Shawna a kol. (2012). V textu je rozebrán vliv dešťů na popularitu politických subjektů v Indii. Podobně může být těmito faktory ovlivněno i volební chování v České republice například znečištění ovzduší a vyšší procentuální zisky stran hlásících se k ochraně životního prostředí.

Vzhledem k charakteru této práce je potřeba zmínit prostorový faktor volebního chování. Podle Bernarda a Kosteleckého (2014) je prostorový faktor chápán jako systém regionálních i globálních faktorů ovlivňující volební rozhodování voličů a politických subjektů na politické scéně. Těmito faktory jsou faktory socioekonomické a fyzicko-demografické vztahující se k danému místu. Vzhledem k této skutečnosti je prostorový aspekt významným vlivem při rozhodování voličů. Konkrétním příkladem může být úspěch hnutí Prísaha na Pohořelicku ve sněmovních volbách roku 2021.

### Prostorové uspořádání volebních výsledků

Prostorovým aspektem, je umožněno identifikovat prostorové vzory volebního chování jednotlivých územních celků. Příkladem budiž volební chování v polských parlamentních a prezidentských volbách. Prostorové chování je podle Zaryckého (2015) ovlivněno například historickou příslušností obyvatelstva daných regionů a ekonomickou aktivitou obyvatelstva. To je patrné z porovnání map volebních výsledků v Polsku a historické příslušnosti územních celků (Obr. 1). Podobná souvislost s historickým územím je patrná také v regionech bývalých Sudet České republiky, kde v současnosti získává hlasy především hnutí ANO 2011 nebo hnutí SPD.



Obr. 1 Srovnání výsledků voleb a historických území v Polsku (The Economist 2018)

## Gerrymandering

Gerrymandering je výrazem pro manipulaci s volebními obvody ve prospěch určitého politického subjektu navzdory historickým hranicím volebního okrsku a je problémem zejména v USA (Duignan, 2024). Tato technika volební manipulace byla poprvé využita Elbridgem Gerrym v roce 1812 ve státě Massachusetts v USA. V rámci překreslení hranic volebních okrsků získal výsledný volební obvod podobu salamandra (mloka). Z tohoto důvodu je od té doby pro tuto techniku zajištění volební většiny využíván termín gerrymandering nebo také gerrymander. V rámci volebního systému České republiky byla určitá forma gerrymanderingu využita stranou ODS v kontextu komunálních voleb v Praze (Blažek 2011). Výzkumu gerrymanderingu je dále věnován článek Ramachandrana a Goldové (2018) zabývající se detekcí odlehklých hodnot při manipulaci s volebními obvody.

## 2.2 Prostorové analýzy volebních výsledků

Prostorové analýzy jsou prováděny z důvodu odhalení vnitřních souvislostí a pochopení hlubšího kontextu volebního chování. V prostorových analýzách volebních výsledků jsou podle Linkeho a O' Loughlina (2016) v současné době řešeny dva hlavní problémy:

- **prostorová autokorelace**, kde je pozorovaná hodnota daného místa geografického prostoru závislá na okolních hodnotách stejné proměnné,
- **Modifiable Areal Unit Problem**, jehož důsledkem je ovlivnění výsledných hodnot současným stavem prostorových jednotek a při změně např. tvaru prostorových jednotek je výsledek odlišný.

### Metoda shlukové analýzy

Ze zahraničních studií na téma volebních výsledků a prostorových analýz byla vybrána práce Nurmiho a Meskanena (2018) využívající shlukovou analýzu pro data finských parlamentních voleb z roku 2010. K výpočtu byl použit speciální algoritmus, jímž byly vymezeny shluky prostřednictvím křivek zobrazujících relativní hodnoty podpory čtyř politických stran. Prostorové analýzy byly v případě analýzy prováděny pro území jihozápadního Finska. Z analýzy dosažených výsledků bylo zjištěno, že nejstabilnější oblastí politického výsledku bylo švédsky hovořící souostroví a rozložení volebních výsledků bylo stejné napříč volebními obvody. Kromě shlukové analýzy byl v rámci práce použit i Herfindahl-Hirschmanův Index (dále jen HHI), jenž je obvykle využíván při analýze trhu. Jedním z nejvýznamnějších subjektů, jímž je HHI aplikován na našem území, je Česká národní banka. Indexem je měřena koncentrace odvětví na trhu, přičemž se hodnota indexu snižuje s nárůstem konkurence více subjektů (Eurostat 2021).

### Metody prostorové autokorelace

Měřením prostorové autokorelace volebních výsledků je umožněno zjistit míru závislosti mezi volebním výsledkem v jedné územní jednotce vzhledem k okolním územním jednotkám.

Jednou z používaných metod k měření prostorové autokorelace je Moranovo I, které je využíváno ke zjištění míry globální autokorelace (Bobbit 2021).

V kontextu analýzy volebních výsledků je výsledek indexu interpretován jako míra shlukování, roztržitosti či náhodného rozmístění podpory politických stran a dalších ukazatelů v rámci celého vymezeného území.

Analýza LISA, která je známa také jako lokální Moranovo I, je (narozdíl od globálního Moranova I nebo Gearyho-C) lokálním indexem prostorové autokorelace. Z pohledu dat volebních výsledků za územní jednotky je umožněno identifikovat místní shluky a odlehle hodnoty zvolených dat volebních výsledků (Anselin 2020a).

V českém prostředí se analýzami prostorové autokorelace zabýval například Kouba (2007), v jehož práci byly použity následující metody:

Prostorový jev	Význam v politice	Statistická metoda
Prostorová autokorelace	Regionalizace stranického systému	Moranovo I kritérium, indikátory LISA
Prostorové režimy	Územní odchylky od celostátního vzorce podpory pro jednotlivé strany	Prostorový regresní model, Chowův test
Prostorová heterogenita	Nacionalizace stranického systému	Giniho Index

Obr. 2 Metody prostorové autokorelace v analýze českého stranického systému

Na základě výsledků analýz prostorové autokorelace bylo konstatováno, že zaznamenaná variabilita podpory jednotlivých politických stran napříč územními celky není podmíněna makroregionálními specifiky. Dalším zajímavým poznatkem je negativní zjištění v odlišnosti prostorových vzorců volebního chování v Čechách a na Moravě. Autorem je v závěru práce nastíněn námět pro budoucí výzkum prostorových procesů v menších územních jednotkách. Výzkum zaměřený na menší územní celky může být přínosem z důvodu detailnějšího pohledu na ovlivnění volebních výsledků.

### Metody hledání odlehlých hodnot

Hledání odlehlých hodnot v rámci datového souboru je proces, při kterém je na základě definované prahové hodnoty diagnostikována odlehlost záznamu vůči zbytku datového souboru (Aggarwal 2016).

Hledání odlehlých hodnot v datech volebních výsledků je zkoumáno v rámci jedné či více proměnných. K odhalení odlehlých hodnot v rámci jedné proměnné jsou používány např. Boxplot nebo Bagplot (Aggarwal 2016). Hodnocení odlehlosti pomocí jednorozměrných metod je velmi efektivní, pokud není potřeba vyhledání odlehlých hodnot u více než jednoho atributu. Pokud je ovšem v rámci šetření vyžadováno vyhledání odlehlých hodnot napříč více atributy, je potřeba využít některou z metod vícerozměrného hodnocení odlehlosti, jakou může být Mahalanobisova vzdálenost (Cansiz, 2021).

K vyhodnocení Brownova a Paynova modelu přechodu voličů využil Mahalanobisovu vzdálenost Forcina a kol. (2011). V rámci modelů byly zjištěny vzdálenosti mezi skutečnými a predikovanými hodnotami mezi evropskými a sněmovními volbami v Itálii. Metodologickému hodnocení odlehlosti z lokálního pohledu je věnován článek Filzmosera a kol. (2014). V článku odlehlost hodnocena na základě definovaného počtu nejbližších sousedů. Tato skutečnost může být přínosem při zkoumání lokálních odlehlostí, kterými mohou být ostrovy v podobě obcí podporující politické strany, jež jsou pro daný region zcela atypické.

### Metody prostorových regresních modelů

V rámci matematických modelů v humánní geografii je využívána celá řada matematických modelů. Mezi tyto modely se řadí také prostorové regresní modely. Hlavním principem prostorových regresních modelů je vysvětlení závisle proměnné na základě jedné či více vysvětlujících proměnných v roli tzv. prediktorů (Rey 2001).

Využití klasických i prostorově vážených regresních modelů v prostředí České republiky je popsáno v článku od Lyska a kol. (2020). Hlavním obsahem článku je popis volebního chování v České republice pomocí SED ukazatelů. Na základě prostorově váženého regresního modelu byl zjištěn odklon tradičně levicového voličstva směrem ke hnutí ANO 2011 zejména v bývalých Sudetech a oblastí vnitřních periferií. Naopak pravicově orientovanými voliči je preferována Občanská demokratická strana (dále jen ODS),



jejímiž hlavními podporovateli jsou především velká i menší města charakterizovaná vyšší mírou podnikatelské činnosti. V porovnání například s politickou situací v USA je toto zjištění překvapivé. Hlavními voliči Demokratické strany jsou obývána především velká města, naopak voliči Republikánské strany žijí především v méně urbanizovaných regionech (Maisel 2016). Změna politické rovnováhy je odrazem celosvětového trendu přesunu části voličů tradičních subjektů, jakými jsou například ODS a ČSSD, k novým politickým subjektům, jakým je například hnutí ANO 2011. V minulosti obdobným subjektem byla také politická strana Věci veřejné (Lysek a kol. 2021).

Dalším příkladem aplikace regresních modelů jsou modely logistické regrese. V kontextu volební geografie je využití logistické regrese popsáno v článku Rusche a kol. (2013), jenž zkoumá data výsledků voleb prezidenta USA v roce 2004. V rámci výzkumu byly aplikovány rozhodovací stromy kombinované s logistickou regresí. Tato metoda je autory článku pojmenována jako LORET. Díky této technice byla predikována účast voliče a podpora určitého politického subjektu. V závěru článku bylo konstatováno, že vytvořený predikční model funguje v porovnání s novými volebními daty lépe na historická volební data. Aplikaci logistické regrese na volební data je dále věnována publikace Rusche a kol. (2013). Hlavním přínosem aplikace logistické regrese na volební data je lepší cílení na určité skupiny voličů jednotlivými politickými stranami prostřednictvím volebních kampaní, programovými prohlášeními apod.

Prostorové analýzy jsou, podle výše uvedených odborných prací, aplikovány v celé řadě domácích i zahraničních studií. Dalším důležitým poznatkem je možnost aplikace metod, jež byly původně určeny k odlišným analýzám, např. HHI. Ze zjištěných skutečností vyplývá vhodnost užitých prostorových analýz, pokud jsou získané informace relevantní.

### **2.3 Odborné práce zabývající se podobným tématem**

Tématu volebních výsledků je věnován velký počet publikací napříč různými obory. Analýze volebních výsledků prostřednictvím odborných článků z politologické oblasti jsou věnovány publikace Mgr. Jakuba Lyska, Ph.D., konzultanta této diplomové práce. Příkladem publikace může být „Continuity in Discontinuity?“ (Lysek a kol. 2022), na kterém se mimo jiné podíleli konzultant a vedoucí této diplomové práce. Podobnými zahraničními publikacemi jsou například publikace od Mészárose a kol. (2007) věnující se analýzám volebních výsledků v Maďarsku či publikace od Zaryckého (2015) provádějící obdobné politologické analýzy polských volebních výsledků. Z českého prostředí je v kontextu analýzy volebních výsledků potřeba zmínit agentury STEM, Median, Kantar apod. Těmito agenturami jsou především vytvářeny modely volebních preferencí, které jsou hlavním ukazatelem aktuální podpory politických subjektů.

Z diplomových prací zabývajících se aplikací prostorových analýz na volební výsledky je příkladem práce Jakuba Vlosinského (2011) z katedry geoinformatiky UPOL na téma prostorové analýzy časových řad volebních výsledků na území Olomouckého kraje. Hlavním cílem práce bylo postihnout případné změny volebních výsledků v průběhu 10 let a vysvětlit jejich důvod, například pomocí korelačních koeficientů.

Dalším příkladem práce s touto tematikou je diplomová práce Jana Ondruška (2012) z katedry politologie MU, jejíž hlavním úkolem bylo zpracování map volebních výsledků v Praze pro roky 2002–2010. V mapách jsou zobrazeny volební výsledky stran ODS, ČSSD, KSČM a KDU–ČSL. Autorem práce jsou také zhodnoceny politické preference Prahy v porovnání s jinými regiony ČR.

Diplomová práce s volební tematikou byla dále zpracována na katedře geografie UPOL Petrem Sixtou (2019). Výsledkem práce byla analýza regionálních rozdílů v rámci

jednotlivých regionů Středočeského kraje a vysvětlení volebního chování na základě SED ukazatelů. Analýzy byly provedeny pro volby, které se uskutečnily mezi lety 1992–2017.

Z výše uvedených odborných publikací vyplývá, že analýza volebních výsledků v prostorovém kontextu je velmi populární tematikou k výzkumu. V rámci prostorových analýz volebních výsledků je uplatňována široká řada metod a postupů, jejichž autory jsou nejen politologové, ale i geografové, matematici a zástupci dalších oborů. S rozvojem informačních technologií a pokročilých výzkumných metod jsou i v oblasti analýzy volebních výsledků stále častěji aplikovány nové nástroje za účelem přesnějšího pochopení a vysvětlení volebních výsledků a jejich predikce do budoucna.

## 3 METODY A POSTUP ZPRACOVÁNÍ

Při zpracování diplomové práce byly využity teoretické metody, jakými jsou například analýza, syntéza a dedukce (Široký 2011). Výběr těchto metod byl podmíněn charakterem práce, jejímž primárním cílem je provést analýzu volebních výsledků. Teoretické metody byly uplatněny na sestavenou datovou sadu z volebních výsledků a SED ukazatelů prostřednictvím prostorových statistických metod.

### 3.1 Použité metody

K analýze dat byly využity především prostorově statistické metody, jejichž obecným principem je práce s daty s informací o umístění v prostoru (Kanade 2022b). Prostorové analýzy lze dále rozdělit na geometrické pracující v rámci definovaného souřadnicového systému (např. dvojrozměrného grafu o souřadnicích X a Y), a geografické obsahující informace o umístění v geografickém prostoru pomocí zeměpisné šířky a délky. Hlavním účelem využití prostorových statistických metod je vysvětlení prostorových vzorů a také snaha o vyřešení nestacionarity zkoumaného jevu, tedy změnu jeho hodnot na poloze v prostoru (Higdon a kol 1999). Konkrétní příklady aplikace prostorových metod jsou uvedeny v podkapitole 4.2.

#### 3.1.1 Mahalanobisova vzdálenost

Mahalanobisova vzdálenost je používána pro vyhledání odlehlých hodnot v rozsahu definovaných atributů datové sady. Mahalanobisova vzdálenost je vzdáleností metrika, kterou je měřena vzdálenost mezi konkrétním bodem a distribucí ostatních bodů. Tuto metriku lze, oproti Euklidovské vzdálenosti, využít ke zjištění vzdálenosti v rámci definovaného rozsahu atributů datové sady (Prabhakaran 2024). Při analýze volebních výsledků byly s pomocí Mahalanobisovy vzdálenosti odhaleny odlehlé hodnoty dat volebních výsledků.

Mahalanobisova vzdálenost  $D^2$  je vypočtena rozdílem  $x$  (vektor odpovídající záznamu v datové sadě o délce definovaných atributů) a  $m$  (vektor nezávislých proměnných). Do výpočtu dále vstupuje proměnná  $C$  (inverzní kovarianční matice nezávislých proměnných). Celý vzorec má následující podobu.

$$D^2 = (x - m)^T \cdot C^{-1} \cdot (x - m) \quad (1)$$

Při analýze odlehlých hodnot byla využita také lokální Mahalanobisova vzdálenost.

Lokální Mahalanobisovou vzdáleností je rovněž měřena vzdálenost mezi více atributy, ale pouze v definovaném okolí záznamů (Rajabzadeh 2018). V rámci diplomové práce byl okruh záznamů definován počtem nejbližších sousedů z geografického hlediska.

#### 3.1.2 Lineární regrese a GWR

Druhou statistickou metodou, která byla v rámci zpracovaných analýz využita, je metoda lineární regrese, jejímž principem je co nejlepší vysvětlení vztahu mezi závislou proměnnou a vysvětlujícími proměnnými, neboli prediktory, s pomocí regresní přímky proložené distribucí vstupních prvků (Moravec 2018). V kontextu diplomové práce byla metoda lineární regrese využita k popisu vztahu mezi závislou proměnnou v podobě procentuálního podílu získaných hlasů politického subjektu a zvolenými SED ukazateli jako prediktory závislé proměnné. Jejím dalším využitím bylo vymezení vhodných prediktorů pro geograficky váženou regresi (dále jen GWR).

Na vzorcí níže je popsána obecná rovnice lineární regrese. Závislá proměnná  $Y$  je vysvětlována hodnotou. Proměnnou  $\beta$  jsou označeny regresní koeficienty udávající míru

změny proměnné  $Y$ , závislou na jednotkové hodnotě změny závislé proměnné, které jsou představovány proměnnou  $X$ . Proměnná  $\varepsilon$  v tomto případě představuje rezidua modelu, která jsou vysvětlena jako odchylka hodnoty od ideálního stavu odpovídajícího umístění na regresní přímce.

$$Y = \beta X + \varepsilon \quad (2)$$

Geograficky vážená regrese je typ regresního modelu, který popisuje vztah mezi závislou proměnnou a vysvětlujícími proměnnými v závislosti na geografickém umístění. Hlavním přínosem je odstranění prostorové autokorelace, což znamená, že určitý prvek v prostoru je závislý na ostatních prvcích v okolí. V rámci GWR je tato skutečnost řešena prostřednictvím vah a přidáním jádrové funkce, která může být fixní (definice na základě vzdálenosti) nebo adaptivní stanovena počtem nejbližších sousedů (Páez 2009). Tvorbou GWR modelů z metodologického hlediska se zabývá Golliniová a kol. (2015) využívající k jejich sestavení balík *GWmodel* z prostředí R. Zmíněný balík poskytuje řadu funkcí k optimalizaci sestaveného GWR modelu, např. *bw.gwr* k výpočtu optimálního vzdálenostního pásma, nebo *gwr.model.selection* pro seřazení atributů podle prostorové významnosti. Z tohoto důvodu byla pro zpracování GWR modelů použita tato metodika. V rámci diplomové práce byly metodou GWR zpracovány čtyři modely vysvětlující prostorovou variabilitu v chování koalice SPOLU, hnutí ANO 2011, koalice Pirátů a Starostů a hnutí SPD jako závislých proměnných a SED ukazatelů jako využitých prediktorů.

### 3.1.3 Logistická regrese

Logistická regrese je statistickou metodou, která je využívána k modelování hodnot závislé proměnné binárního charakteru. Podobně jako u lineární regrese jsou i v případě logistické regrese k vysvětlení použity prediktory. Distribuce hodnot ve vytvořeném modelu je namísto přímky popsána sigmoidální křivkou. V kontextu logistické regrese jsou používány pojmy pravděpodobnost a šance. Mírou pravděpodobnosti je udávána možnost výskytu a absence binárního jevu. Pojem šance je pak definován poměr pravděpodobnosti výskytu a absence daného jevu (Kanade 2022). V diplomové práci byla porovnána míra úspěšnosti predikce na základě klasických a prostorově vážených modelů logistické regrese.

### 3.1.4 Shluková analýza

Shluková analýza je statistická metoda, kterou jsou na základě určité metody či algoritmu prvky zařazeny do shluků společně s prvky podobných vlastností. Podle cíle shlukové analýzy lze podle Šarmanové (2012) rozdělit metody na nehierarchické, jejichž cílem je rozklad prvků na podmnožiny, a hierarchické, kdy každý nově vytvořený shluk je zjemněním předcházejícího shluku. Výhodou hierarchického shlukování oproti nehierarchickému je snadnější interpretace a odvozování nových shluků s pomocí dendrogramu. Důvodem je, že vytvoření nového počtu shluků u nehierarchického shlukování je podmíněno novým výpočtem shluků, kdy nově vytvořené shluky nemusí mít žádnou souvislost s předchozími shluky.

V rámci tvorby diplomové práce bylo ke zpracování vybráno hierarchické aglomerativní shlukování. Hlavním výhodou aglomerativního shlukování je oproti diviznímu přístupu způsob tvoření shluků. Tvorba shluků je započata v jednotlivých prvcích datového souboru, které jsou postupně spojovány do větších shluků (Šarmanová 2012). Hierarchické aglomerativní shlukování bylo využito k vymezení shluků na základě volebních výsledků sedmi nejúspěšnějších subjektů ve sněmovních volbách 2021.

### **3.1.5 Metody výpočtu volebních ukazatelů**

Stažená data volebních výsledků jsou v datové sadě vedena jako zápisy absolutních hodnot. Z tohoto důvodu bylo nutné provést přepočtení dat na ukazatele volebních úspěchů jednotlivých politických subjektů do procentuální podoby. K přepočtení dat byla využita metodika ČSÚ (2015), podle které je procentuální volební úspěch vypočten jako podíl platných hlasů pro daný politický subjekt na celkovém počtu platných hlasů.

Dále byla vypočtena i volební účast pro sněmovní volby 2021 a prezidentské volby 2023 jako podíl počtu vydaných obálek na celkovém počtu voličů uvedených v seznamu pro dané volby (ČSÚ 2015).

## **3.2 Použitá data**

Diplomová práce je zaměřena na prostorové hodnocení dat volebních výsledků za použití SED ukazatelů. Z tohoto důvodu byla nejprve shromážděna data volebních výsledků. Všechna data byla shromážděna za rok 2021 z důvodu jednotnosti datové sady. Jedinou výjimkou byl index kvality života a data z voleb prezidenta republiky, která byla shromážděna za rok 2023.

### **3.2.1 Data volebních výsledků**

Data volebních výsledků jsou složena z volně dostupných dat, která byla stažena z datového portálu ČSÚ za roky 2021 a 2023 ve formátu XLSX. Shromážděná data volebních výsledků jsou pro každé volby složena z tabulky volebních výsledků jednotlivých kandidujících subjektů a tabulky s dalšími statistickými charakteristikami, např. počet platných hlasů, počet voličů v seznamu, počet odevzdaných obálek podle druhu nebo počet platných hlasů. Atributy související s obálkami popisují počet odevzdaných obálek, který odpovídá počtu vhozených obálek do volební urny, nebo počet vydaných obálek, kdy není podmíněno její vhození do volební urny. Je však nutné podotknout, že oba ukazatele jsou ve většině případů shodné. Posledním volebním ukazatelem jsou platné hlasy odpovídající součtu platných hlasujících lístků v obálkách. Neplatnými hlasy jsou označeny například vážně poškozené, či jinak znehodnocené volební lístky. Stažená data odpovídala podrobnosti volebních okrsků, která jsou na stránkách ČSÚ rovněž k dispozici ke stažení ve formátu SHP. Data volebních výsledků jsou primárním objektem pro stanovení cílů jednotlivých analýz.

### **3.2.2 Data SED ukazatelů**

Následujícím krokem bylo shromáždění dat SED ukazatelů. Tato část datové sady byla získána prostřednictvím geoportálu Sčítání 2021, který vznikl jako jeden z výstupů SLBD 2021. Jednalo se například o data rozmístění a hustoty obyvatel, ekonomické aktivity nebo vzdělání obyvatelstva. Data byla stažena ve formátu CSV. SED ukazatele byly dále obohaceny o data, která byla využita v rámci výpočtu indexu sociálního vyloučení. Tento výzkum byl proveden Agenturou pro sociální začleňování (2021), která je součástí Ministerstva práce a sociálních věcí (dále jen MPSV). Příkladem přidaných dat budiž podíl osob v exekucích nebo podíl osob pobírajících příspěvek na bydlení. Data byla obdržena od vedoucího práce ve formátu SHP.

Jedním ze dvou dalších přidaných ukazatelů byl index kvality života, jehož metodika je popsána v článku „Implementation of GIS Tools in the Quality of Life Assessment of Czech Municipalities“ (Macků a kol. 2023). Druhým přidaným ukazatelem byl index příslušnosti k městu nebo venkovu, jehož metodika byla popsána v článku „Using a fuzzy

inference system to delimit rural and urban municipalities in the Czech republic in 2010“ (Pászto a kol. 2010). Data těchto indexů opět poskytl vedoucí diplomové práce ve formátu SHP.

### **3.3 Použité programy**

Pro provedení statistických prostorových analýz a zpracování výstupů byly využity níže popsané programy.

#### **3.3.1 ArcGIS Pro**

V rámci programu ArcGIS Pro verze 3.1.3 byly především zpracovány mapové výstupy jednotlivých prostorových statistických analýz. Program byl dále využit pro vytvoření prostorově váženého logistického modelu.

#### **3.3.2 QGIS**

Program QGIS verze 3.28.3 byl využit pro úpravu a převod mezi prostorovými formáty OGC Geopackage (dále jen GPKG) a Esri Shapefile (dále jen SHP). Úprava dat a následná konverze byla provedena při zpracování výstupu GWR modelů.

#### **3.3.3 MS Excel**

Pro vytvoření doprovodných statistických materiálů, jakými byly grafy, tabulky apod., byl využit program Microsoft Excel.

#### **3.3.4 RStudio**

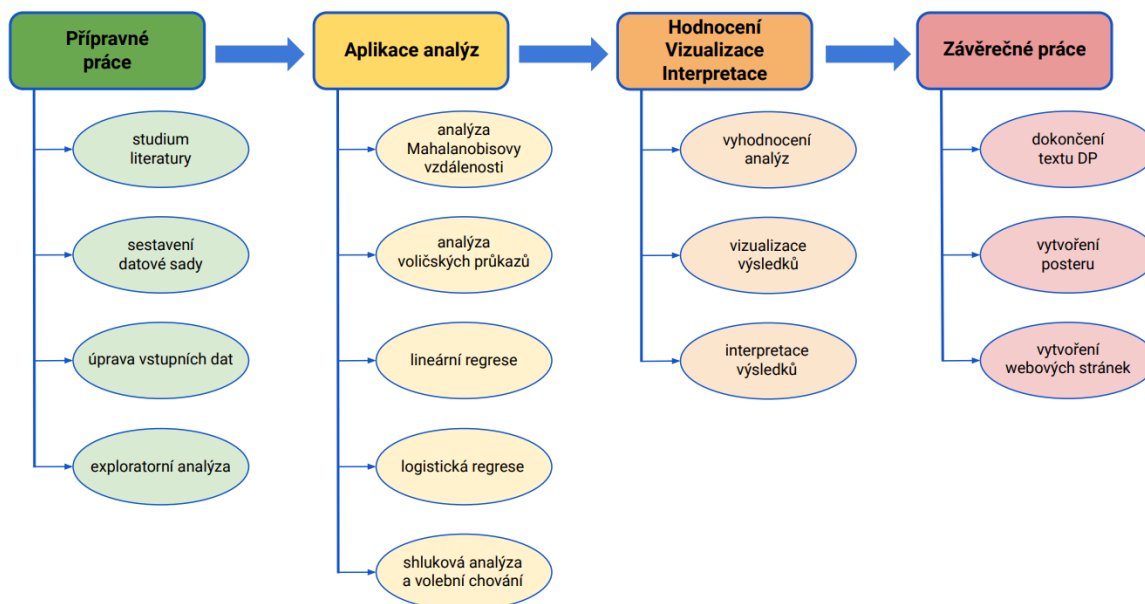
Program RStudio je vývojovým prostředím pro programovací a statistický jazyk R. Pro vytvoření výstupů prostorových statistických analýz byla použita verze jazyka R 4.2.2.

#### **3.3.5 Inkscape**

Tvorba schémat, úprava mapových výstupů a další vizualizace byly provedeny prostřednictvím grafického programu Inkscape verze 1.3.2.

### 3.4 Postup zpracování

Postup zpracování je vizualizován formou vývojového diagramu (Obr. 3). Jednotlivé kroky zpracování samotné diplomové práce jsou popsány v podkapitolách 4.1 a 4.2. Skutečný postup vypracování byl oproti diagramu pracovního postupu měněn pouze minimálně.



Obr. 3 Postup zpracování diplomové práce

## 4 VLASTNÍ ŘEŠENÍ

V této kapitole je sepsán postup praktické části diplomové práce. První část je věnovaná procesu sestavení datové sady a zisku základních poznatků prostřednictvím exploratorní analýzy. V druhé části je popsán proces aplikace prostorových analýz za účelem detekce odlehklých hodnot, popisu prostorové variability volebních výsledků prostřednictvím GWR modelů, vymezení shluků volebního chování apod.

### 4.1 Sestavení datové sady

Po získání teoretických základů prostřednictvím literární rešerše bylo přistoupeno k sestavení datové sady z dat volebních výsledků a SED ukazatelů. Po získání dat, jejichž zdroje jsou popsány v podkapitole 3.2, bylo zjištěno, že pro provedení samotných prostorových analýz bude nejdříve zapotřebí data volebních výsledků upravit.

#### 4.1.1 Data volebních výsledků

Prvním problémem byl způsob zapsaných dat v tabulce výsledků sněmovních voleb 2021, kdy počet záznamů v datech převyšoval počet volebních okrsků. Příčinou této skutečnosti bylo zaznamenání preferenčních hlasů kandidátů politických subjektů v tabulce. Z tohoto důvodu byly v jednotlivých sloupcích tabulky zaznamenány preferenční hlasy kandidátů, namísto politických subjektů samotných. Počet záznamů byl určen podle počtu politických stran a hnutí, které v daném okrsku obdržely alespoň jeden hlas. Dále bylo zjištěno, že jednotlivé kandidující politické subjekty jsou rozlišeny formou čísel v atributu **KSTRANA**, která byla shodná s přiřazeným číslem ve sněmovních volbách 2021.

ID_OKRSKY	TYP_FORM	OPRAVA	CHYBA	OKRES	OBEC	OKRSEK	KC_1	KSTRANA	POC_HLASU	KC_2	HLASY_01	HLASY_02	HLASY_03	HLASY_04	HLASY_05
7	2	0	0	1100	500054	1000	2	1	8	9	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	2	1	3	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	3	7	10	0	1	0	2	0
7	2	0	0	1100	500054	1000	2	4	27	31	1	2	1	2	0
7	2	0	0	1100	500054	1000	2	5	9	14	0	1	0	0	0
7	2	0	0	1100	500054	1000	2	7	2	9	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	8	20	28	1	2	3	3	1
7	2	0	0	1100	500054	1000	2	10	1	11	1	0	1	0	0
7	2	0	0	1100	500054	1000	2	12	6	18	0	0	1	1	0
7	2	0	0	1100	500054	1000	2	13	216	229	35	16	12	16	15
7	2	0	0	1100	500054	1000	2	15	1	16	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	16	2	18	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	17	143	160	25	4	20	5	5
7	2	0	0	1100	500054	1000	2	18	7	25	0	0	0	0	0
7	2	0	0	1100	500054	1000	2	20	21	41	0	0	0	0	0
8	2	0	0	1100	500054	1001	3	1	14	15	2	3	1	0	1
8	2	0	0	1100	500054	1001	3	2	2	4	0	0	1	0	0
8	2	0	0	1100	500054	1001	3	3	7	10	3	2	2	0	0
8	2	0	0	1100	500054	1001	3	4	20	24	1	0	0	5	0
8	2	0	0	1100	500054	1001	3	5	8	13	2	2	0	1	0
8	2	0	0	1100	500054	1001	3	8	13	21	2	0	0	1	0
8	2	0	0	1100	500054	1001	3	11	1	12	0	0	0	0	0
8	2	0	0	1100	500054	1001	3	12	4	16	0	0	0	0	0
8	2	0	0	1100	500054	1001	3	13	232	245	51	46	35	36	31
8	2	0	0	1100	500054	1001	3	15	1	16	1	0	0	0	0
8	2	0	0	1100	500054	1001	3	17	145	162	39	17	31	7	8
8	2	0	0	1100	500054	1001	3	18	6	24	0	0	0	0	0
8	2	0	0	1100	500054	1001	3	20	57	77	4	1	1	1	0
9	2	0	0	1100	500054	1002	4	1	8	9	0	1	1	0	1
9	2	0	0	1100	500054	1002	4	2	1	3	0	0	0	0	0
9	2	0	0	1100	500054	1002	4	3	6	9	2	1	1	0	0
9	2	0	0	1100	500054	1002	4	4	26	30	0	1	0	5	0
9	2	0	0	1100	500054	1002	4	5	21	26	4	1	0	1	0
9	2	0	0	1100	500054	1002	4	8	14	22	1	0	3	1	0
9	2	0	0	1100	500054	1002	4	12	13	25	2	1	0	1	1

Obr. 4 Ukázka vstupních dat pro sněmovní volby 2021

Z tohoto důvodu byly nejdříve provedeny součty preferenčních hlasů pro kandidující subjekty. Dále byla čísla subjektů ve sloupci **KSTRANA** nahrazena zkratkou, pod kterou budou v sestavené datové sadě vedena. Nakonec byl vytvořen jednoduchý skript pro seskupení dat podle identifikátoru volebního okrsku:



```

library(dplyr)
grouping<-db %>%
group_by(okrsek) %>%
summarise('SZE_1' = sum(SZE_1),
          'SVYDE_2' = sum(SVYDE_2),
          'VOBL_3' = sum(VOBL_3),
          'SPD_4' = sum(SPD_4),
          'CSSD_5' = sum(CSSD_5),
          'PRABLO_6' = sum(PRABLO_6),
          'ALNASI_7' = sum(ALNASI_7),
          'TSS_8' = sum(TSS_8),
          'ALPROBU_9' = sum(ALPROBU_9),
          'PRAM_10' = sum(PRAM_10),
          'LEV_11' = sum(LEV_11),
          'PRI_12' = sum(PRI_12),
          'SPOLU_13' = sum(SPOLU_13),
          'SEN_14' = sum(SEN_14),
          'NEV_15' = sum(NEV_15),
          'KOCE_16' = sum(KOCE_16),
          'PIRSTAN_17' = sum(PIRSTAN_17),
          'KSCM_18' = sum(KSCM_18),
          'MOZEHN_19' = sum(MOZEHN_19),
          'ANO_20' = sum(ANO_20),
          'OTCE_21' = sum(OTCE_21),
          'MOR_22' = sum(MOR_22)
)

```

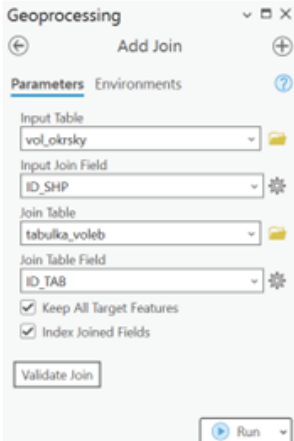
Obr. 5 Ukázka skriptu pro seskupení kandidujících subjektů

Data z prezidentských voleb 2023 byla složena ze dvou tabulek pro první i druhé kolo. Samotná data byla v tomto případě přehledně členěna do sloupců, kdy byl každý jednotlivý sloupec přiřazen určitému kandidátovi podle přiřazeného čísla podobně, jako v případě sněmovních voleb 2021. Z tohoto důvodu nebylo zapotřebí aplikovat stejný postup a jedinou úpravou bylo v tomto případě spojení obou tabulek volebních výsledků a přejmenování atributů.

Poté, co byla data připravena, byla spojena s prostorovou vrstvou volebních okrsků pomocí nástroje *Add join* v ArcGIS Pro. Za účelem vytvoření univerzálního identifikátoru pro každý okrsek byla spojena pole **OBEC** a **ID\_OKRSKY** v tabulce volebních výsledků. Je potřeba zde zmínit, že obec v kontextu těchto dat odpovídá ZUJ, neboť v prostředí velkých měst, například Prahy, je území členěno na městské části, a nikoliv jako jedna obec, jak je tomu například v polygonové vrstvě obcí z databáze ArcČR 500. V prostorové vrstvě volebních okrsků ve formátu SHP byly se stejným záměrem spojeny sloupce **Momc**, který odpovídá městským částem a **Cislo** odpovídající číslování okrsků.

OBEC	ID_OKRSKY	ID_TAB
500054	7	5000547
500054	8	5000548
500054	9	5000549
500054	10	50005410
500054	11	50005411
500054	12	50005412
500054	13	50005413
500054	14	50005414
500054	15	50005415
500054	16	50005416
500054	17	50005417
500054	18	50005418
500054	19	50005419
500054	20	50005420
500054	21	50005421

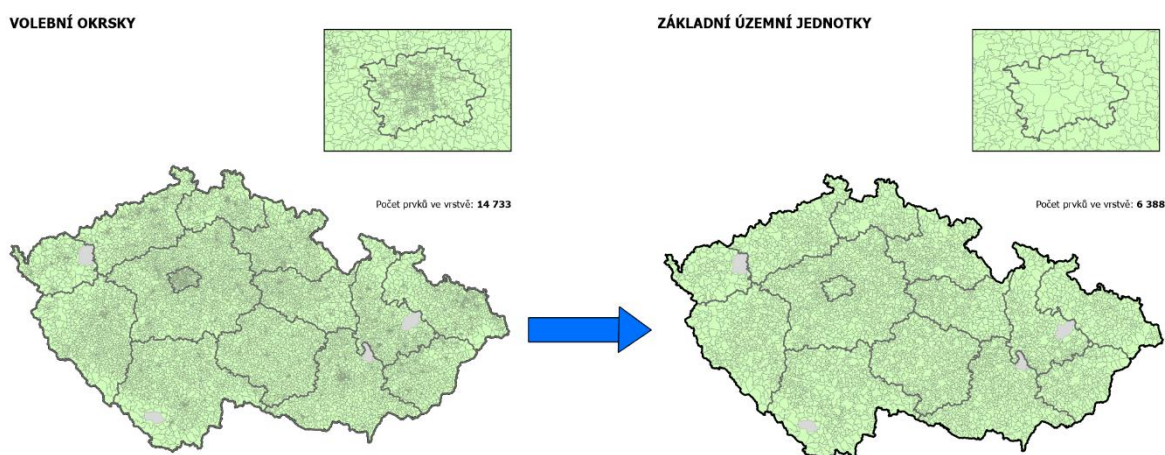
Momc	Cislo	ID_SHP
500054	7	5000547
500054	8	5000548
500054	9	5000549
500054	10	50005410
500054	11	50005411
500054	12	50005412
500054	13	50005413
500054	14	50005414



Obr. 6 Spojení dat prostřednictvím vytvořených identifikátorů v nástroji Add Join

### **Agregace dat volebních okrsků na ZUJ a přepoččet hodnot**

Při shromažďování dat SED ukazatelů bylo zjištěno, že maximální mírou podrobnosti jsou základní územní jednotky (dále jen ZUJ). Po konzultaci s politologem Mgr. Jakubem Lyskem, Ph.D., který je konzultantem této diplomové práce, bylo zjištěno, že příslušná data jsou vytvářena pouze prostřednictvím zadání zakázky ČSÚ. Z tohoto důvodu nejsou volně dostupné SED ukazatele s rozpisem za volební okrsky. Podobná data byla obdržena v podobě prostorové vrstvy ve formátu SHP při zadání diplomové práce, kdy jsou v podrobnosti okrsků zaznamenány volební výsledky sněmovních voleb za rok 2010 a data ze SLBD 2011. Vzhledem k zastaralosti těchto dat proto bylo přistoupeno k agregaci dat volebních výsledků na ZUJ a následnému provedení statistických analýz s touto prostorovou podrobností. K samotné agregaci volebních dat na ZUJ byl využit již identifikátor městských částí **Momc**.



*Obr. 7 Převod prostorové podrobnosti z volebních okrsků na ZUJ*

Dalším krokem po agregaci bylo přepočtení volebních výsledků z podoby absolutního počtu hlasů jednotlivých politických subjektů na procentuální podíly pro jednotlivé politické subjekty. Nakonec byla vypočtena volební účast pro jednotlivá ZUJ v rámci sněmovních voleb 2021 a prezidentských voleb 2023. Způsob výpočtu procentuálních podílů a volební účasti je popsán v podkapitole 3.1.5.

#### **4.1.2 Data SED ukazatelů**

Po konverzi dat na prostorovou podrobnost SED ukazatelů bylo přistoupeno k jejich přidání do datové sady. Vzhledem k povaze stažených dat z geoportálu SLBD 2021, která již byla přepočtena na procentuální hodnoty a obsahovala i identifikátor pro ZUJ, nebyly zapotřebí žádné speciální kroky pro spojení s daty volebních okrsků.

V případě dat využitých pro výpočet indexů kvality života, sociálního vyloučení a města a venkova, včetně indexů samotných, byla provedena disagregace, neboť data byla obdržena s prostorovou podrobností odpovídající obcím. Ačkoliv tato data nebyla stejné prostorové podrobnosti, bylo po konzultaci s politologem a vedoucím práce rozhodnuto o jejich ponechání. Důvodem byla především tematická odlišnost dat v porovnání s daty ze SLBD 2021, která spočívala kupříkladu v komplexním hodnocení kvality života nebo sociálního vyloučení. Proto byla provedena substituce jednotnou hodnotou odpovídající průměru obce s dalším členěním na městské části.

Tab. 1 Sestavená datová sada složená z volebních výsledků SED ukazatelů

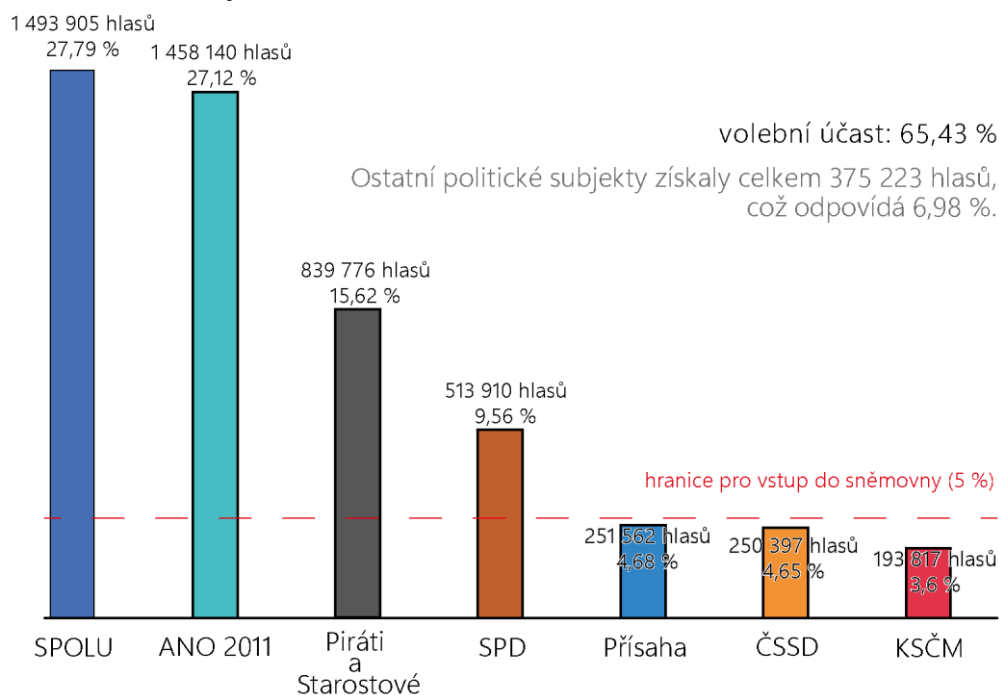
SESTAVENÁ DATOVÁ SADA		
Zkratka	Vysvětlivka	Podrobnost dat
SZE	strana Zelených	základní územní jednotky
SVYDE	Švýcarská demokracie	základní územní jednotky
VOBL	Volný Blok	základní územní jednotky
SPD	Svoboda a přímá demokracie	základní územní jednotky
CSSD	Česká strana sociálně demokratická	základní územní jednotky
PRABLO	Volte Pravý Blok www.cibulka.net	základní územní jednotky
ALNASI	Aliance národních sil	základní územní jednotky
TSS	Trikolora Svobodní Soukromníci	základní územní jednotky
ALPROBU	Aliance pro budoucnost	základní územní jednotky
PRAM	Hnutí Prameny	základní územní jednotky
LEV	strana Levice	základní územní jednotky
PŘI	hnutí Přísaha	základní územní jednotky
SPOLU	koalice SPOLU	základní územní jednotky
SEN	Senioři 21	základní územní jednotky
NEV	Nevolte Urza.cz	základní územní jednotky
KOCE	Koruna Česká	základní územní jednotky
PIRSTAN	koalice Pirátů a Starostů	základní územní jednotky
KSCM	Komunistická strana Čech a Moravy	základní územní jednotky
MOZEHN	Moravské zemské hnutí	základní územní jednotky
ANO	hnutí ANO 2011	základní územní jednotky
OTCE	Otevřeme Česko	základní územní jednotky
MOR	strana Moravané	základní územní jednotky
VOL_UCAST_SN	volební účast ve sněmovních volbách	základní územní jednotky
PF_01	Pavel Fisher – první kolo prezidentských voleb	základní územní jednotky
JB_01	Jaroslav Bašta – první kolo prezidentských voleb	základní územní jednotky
PP_01	Petr Pavel – první kolo prezidentských voleb	základní územní jednotky
TZ_01	Tomáš Zima – první kolo prezidentských voleb	základní územní jednotky
DN_01	Danuše Nerudová – první kolo prezidentských voleb	základní územní jednotky
AB_01	Andrej Babiš – první kolo prezidentských voleb	základní územní jednotky
KD_01	Karel Diviš – první kolo prezidentských voleb	základní územní jednotky
MH_01	Marek Hilšer – první kolo prezidentských voleb	základní územní jednotky
VOL_UCAST_PR_01	volební účast v prvním kole prezidentských voleb	základní územní jednotky
PP_02	Petr Pavel – druhé kolo prezidentských voleb	základní územní jednotky
AB_02	Andrej Babiš – druhé kolo prezidentských voleb	základní územní jednotky
VOL_UCAST_PR_02	volební účast ve druhém kole prezidentských voleb	základní územní jednotky
prum_vek	průměrný věk obyvatelstva	základní územní jednotky
in_star	index stáří	základní územní jednotky
mira_nez	míra nezaměstnanosti	základní územní jednotky
po_zam_os	podíl zaměstnaných osob	základní územní jednotky
prim_s	podíl obyvatel pracujících v primárním sektoru	základní územní jednotky
sek_s	podíl obyvatel pracujících v sekundárním sektoru	základní územní jednotky
ter_kv_s	podíl obyvatel pracujících v terciárním a kvartérním sektoru	základní územní jednotky
po_svob	podíl svobodných osob	základní územní jednotky
po_ze_vda	podíl ženatých nebo vdaných osob	základní územní jednotky
po_rozv	podíl rozvedených osob	základní územní jednotky
po_ovd	podíl ovdovělých osob	základní územní jednotky
po_byd_ob	podíl obyvatel s bydlištěm v obci trvalého pobytu	základní územní jednotky
po_byd_okr	podíl obyvatel s bydlištěm v jiné části stejného okresu jako trvalé bydliště	základní územní jednotky
po_byd_o_k	podíl obyvatel s bydlištěm v jiné části okresu stejného kraje jako trvalé bydliště	základní územní jednotky
po_byd_kr	podíl obyvatel s bydlištěm v jiném kraji než kraj trvalého bydliště	základní územní jednotky
po_byd_zah	podíl obyvatel s bydlištěm v zahraničí	základní územní jednotky
po_z_s_0_d	podíl žen bez živě narozených dětí	základní územní jednotky
po_z_s_1_d	podíl žen s 1 živě narozeným dítětem	základní územní jednotky
po_z_s_2_d	podíl žen s 2 živě narozenými dětmi	základní územní jednotky
po_z_s_3	podíl žen s 3 a více živě narozenými dětmi	základní územní jednotky
po_bez_vi	podíl obyvatelstva bez náboženské víry	základní územní jednotky
po_csh	podíl obyvatelstva hlásících se k církvi československé husitské	základní územní jednotky
po_ci_ev	podíl obyvatelstva hlásících se k církvi české evangelické církvi	základní územní jednotky
po_rim_kat	podíl obyvatelstva hlásících se k římskokatolické církvi	základní územní jednotky
po_neuv	podíl obyvatelstva s neuvedenou vírou	základní územní jednotky
po_prav	podíl obyvatelstva hlásících se k pravoslaví	základní územní jednotky
po_v_b_n_p	podíl obyvatelstva které je věřící ale nehlásí se k žádné církvi	základní územní jednotky
z_vz	základní vzdělání	základní územní jednotky
vz_str_b_m	střední vzdělání bez maturitní zkoušky	základní územní jednotky
vz_str_s_m	střední vzdělání s maturitní zkouškou	základní územní jednotky
vs_vz	vysokoškolské vzdělání	základní územní jednotky
poc_ob	počet obyvatel	základní územní jednotky
po_ob_exe	podíl obyvatel v exekčním řízení	obce
po_pri_by	podíl obyvatel s příspěvkem na bydlení	obce
po_ucha_UP	podíl obyvatel registrovaných na úřadu práce	obce
po_p_uk_zs	podíl obyvatel s předčasně ukončeným základním vzděláním	obce
in_soc_vyl	index sociálního vyloučení	obce
in_me_ve	index města a venkova	obce
in_kv_ziv	index kvality života	obce

### 4.1.3 Exploratorní analýza

Dalším krokem po sestavení datové sady byla exploratorní analýza za účelem získání základních poznatků o povaze sestavených dat a navržení hypotéz pro prostorové analýzy. Zkoumaná datová sada je složena ze 36 ukazatelů volebních výsledků a 39 SED ukazatelů. Datová sada byla vytvořena v prostorové i tabulkové verzi s rozsahem 6 388 záznamů. Popis operací provedených prostřednictvím skriptu exploratorní analýzy v jazyce R je součástí přílohy.

#### Přehled volebních výsledků politických subjektů

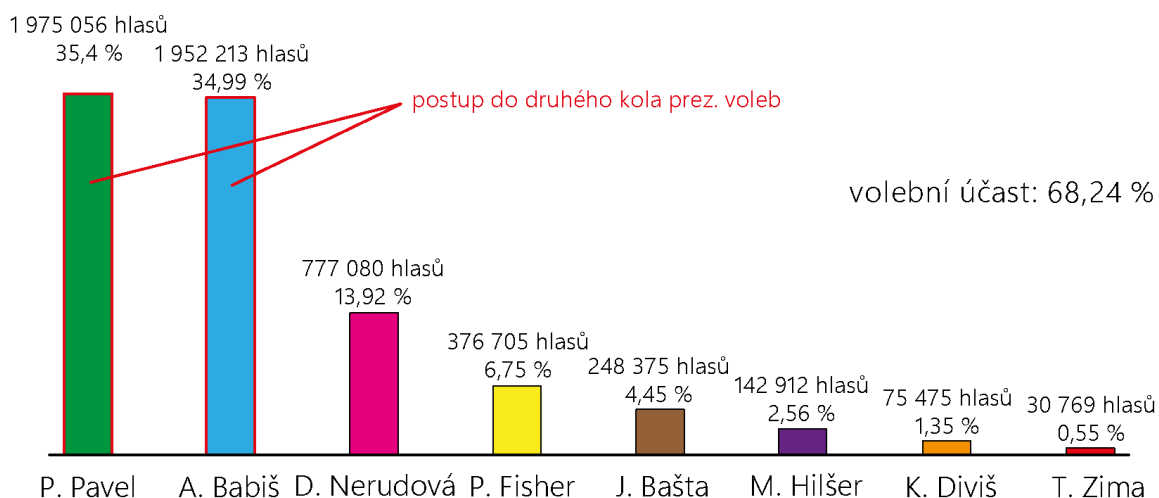
Prvním krokem exploratorní analýzy bylo vytvoření statistických přehledů analyzovaných voleb. Prvními analyzovanými volbami byly sněmovní volby 2021, kde zvítězila koalice SPOLU s 1 493 905 hlasy. Hnutí ANO 2011 skončilo na druhém místě s 27,12 % všech odevzdaných hlasů. Podmínky pro vstup do poslanecké sněmovny byl dále splněn koalici Pirátů a Starostů a hnutím SPD se značným odstupem od prvního a druhého místa. Zajímavou skutečností je přes 1 000 000 propadlých hlasů, jež byla nejvíce patrná u hnutí Přísaha a strany ČSSD, kterými jen těsně nebyly splněny podmínky pro vstup do poslanecké sněmovny Parlamentu ČR.



Obr. 8 Přehled volebních výsledků sněmovních voleb 2021

(zdroj dat: ČSÚ)

Prezidentské volby 2023 jsou byly druhými analyzovanými volbami. V prvním kole prezidentských voleb zvítězil občanský kandidát Petr Pavel, který v minulosti zastával vysoké pozice ve strukturách NATO i Armády České republiky. Druhý v pořadí skončil bývalý premiér ČR a předseda hnutí ANO 2011 Andrej Babiš. Oba tyto kandidáti se následně utkali ve druhém kole voleb hlavy státu. Třetí v pořadí se umístila Danuše Nerudová působící v minulosti jako rektorka Mendelovy univerzity v Brně. Následoval senátor za stranu KDU–ČSL Pavel Fisher, jež získal necelých 7 % hlasů. Dalšími účastníky prvního kola prezidentských voleb byli Jaroslav Bašta, Marek Hilšer, Karel Diviš a Tomáš Zima. Volební účast prezidentských voleb byla v porovnání se sněmovními volbami 2021 vyšší o 2,81 % hlasů.



Obr. 9 Přehled volebních výsledků prvního kola prezidentských voleb 2023

(zdroj dat: ČSÚ)

Ve druhém kole prezidentských voleb opět zvítězil Petr Pavel se ziskem 3 359 151 hlasů a obhájil tak výhru z prvního kola nad Andrejem Babišem. Volební účast druhého kola se zvýšila na 70,25 %, čímž byla překonána účast v prvním kole prezidentských voleb (ČSÚ 2023).

### Pohled na nejúspěšnější politické subjekty

Po sestavení grafů obecného přehledu volebních výsledků sněmovních voleb 2021 a prezidentských voleb 2023 bylo přistoupeno k vizualizaci jejich prostorového rozložení. Za tímto účelem byly sestaveny dvě mapové kompozice.

V mapách sněmovních voleb je při pohledu na dva nejúspěšnější politické subjekty, kterými byla koalice SPOLU a hnutí ANO 2011, patrné několik klíčových charakteristik. První charakteristikou je obecný úspěch koalice SPOLU ve dvou nejlidnatějších městech ČR, tedy v Praze a v Brně, a v jejich nejbližším okolí. Dále byl získán vyšší podíl hlasů také ve většině krajských měst, s výjimkou Ostravy a Ústí nad Labem. Velmi slabé výsledky byly naopak zaznamenány v Moravskoslezském a Ústeckém kraji.

Při pohledu na volební výsledky hnutí ANO 2011 a porovnání s předchozí mapou bylo zjištěno, že vyšší podíl hlasů byl získáván právě v regionech neúspěchu koalice SPOLU. Nejsilnější volební výsledek byl obdrženo na Ústecku a Ostravsku, naopak velmi slabý v okolí Prahy.

Koalice Pirátů a Starostů zaznamenala úspěch v obdobných regionech jako koalice SPOLU. Hlavním rozdílem je úspěch také v okresech Kolín, Nymburk a Mladá Boleslav, které tvoří souvislý pás končící na Liberecku. Koalice Pirátů a Starostů byla v porovnání s vítězem voleb méně úspěšná na jihu Moravy a na českomoravském pomezí.

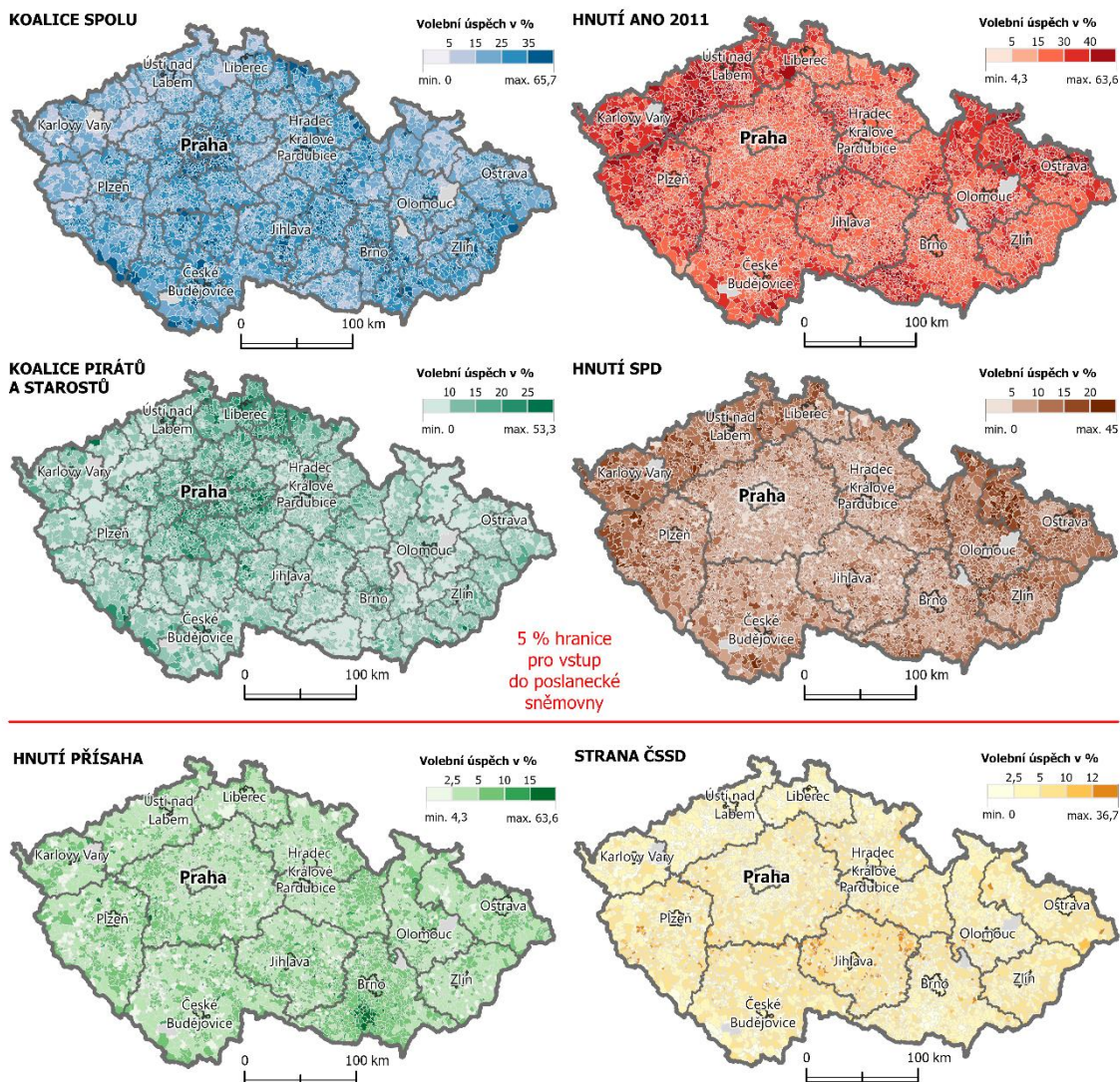
Hnutí SPD bylo posledním politickým subjektem, který byl na základě volebních výsledků přijat do poslanecké sněmovny. Nejvyšší procentuální podíl je patrný v podobných regionech jako u hnutí ANO 2011. Při bližším zkoumání je však patrná vyšší míra rozptýlenosti u ZUJ s vysokým podílem procentuálních hlasů.

Hnutí Přísaha a strana ČSSD jsou dalšími politickými subjekty, jimiž však nebyla překonána hranice 5 % pro vstup do poslanecké sněmovny. Hnutí Přísaha byla nejúspěšnější na Pohořelicku. Přímo v Pohořelicích se 36% podílem hlasů bylo hnutí Přísaha vítězné, v kontextu celé ČR byl tento výsledek zcela unikátní.

V porovnání strany ČSSD a hnutí Prísaha nebyl u ČSSD zjištěn žádný podobně úspěšný region. Obce s výraznějším úspěchem jsou rozmístěny nahodile a jediným více úspěšným regionem je západ kraje Vysočina.

### NEJÚSPĚŠNĚJŠÍ STRANY HNUTÍ A KOALICE SNĚMOVNÍCH VOLEB

v České republice v roce 2021

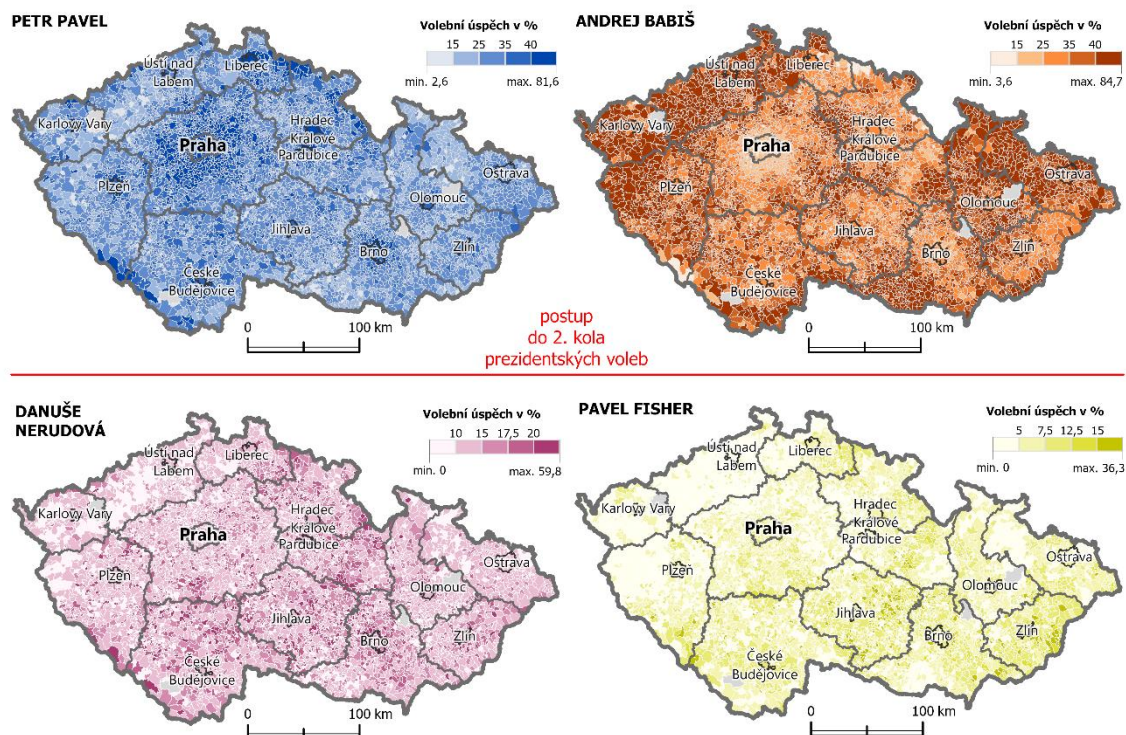


Obr. 10 Volební zisky nejúspěšnějších subjektů ve sněmovních volbách 2021

Volební výsledek prezidentských voleb u Petra Pavla a Andreje Babiše se podobá volebnímu výsledku koalice SPOLU a hnutí ANO 2011. Petr Pavel byl opět vítězem voleb ve většině krajských měst, naopak menší úspěch byl patrný v okrajových regionech, kde byl úspěšnější Andrej Babiš.

Danuše Nerudová a Pavel Fisher byly dalšími kandidáty s nejvyšším volebním ziskem. Danuše Nerudová byla podpořena vyšším volebním ziskem na českomoravském pomezí, ale také v Brně a v jižních Čechách. Hlavními oblastmi podpory Pavla Fishera byly jižní Morava a jižní Čechy.

## NEJÚSPĚŠNĚJŠÍ KANDIDÁTI PREZIDENTSKÝCH VOLEB v České republice v roce 2023



Obr. 11 Volební zisky nejúspěšnějších subjektů v prezidentských volbách 2023

Z výše uvedených zjištění bylo vyvozeno několik závěrů. Prvním je vyšší podpora nynější vládní koalice v regionech, které jsou obecně chápány jako bohatší a rozvinutější. Naopak vysoká podpora opozičních stran je opřena zejména o okrajové regiony, například Bruntálsko a Ústecko. Tato charakteristika může být aplikována také na prezidentské volby, kdy Petra Pavla a Danuši Nerudou podpořily spíše oblasti, kterými byly ve sněmovních volbách 2021 voleny spíše subjekty vládní koalice; a v oblastech s nižším volebním ziskem těchto kandidátů je obvykle vykazován významný volební zisk Andreje Babiše.

Dalším zajímavým zjištěním je úspěch Pavla Fishera v regionech, které jsou charakteristické vyšším podílem věřících hlásících se k římskokatolické církvi v porovnání se zbytkem ČR. Tento prostorový vzor je více patrný v prezidentských volbách, neboť ve sněmovních volbách strana KDU–ČSL opírající se zejména o tyto regiony kandidovala v rámci koalice SPOLU.

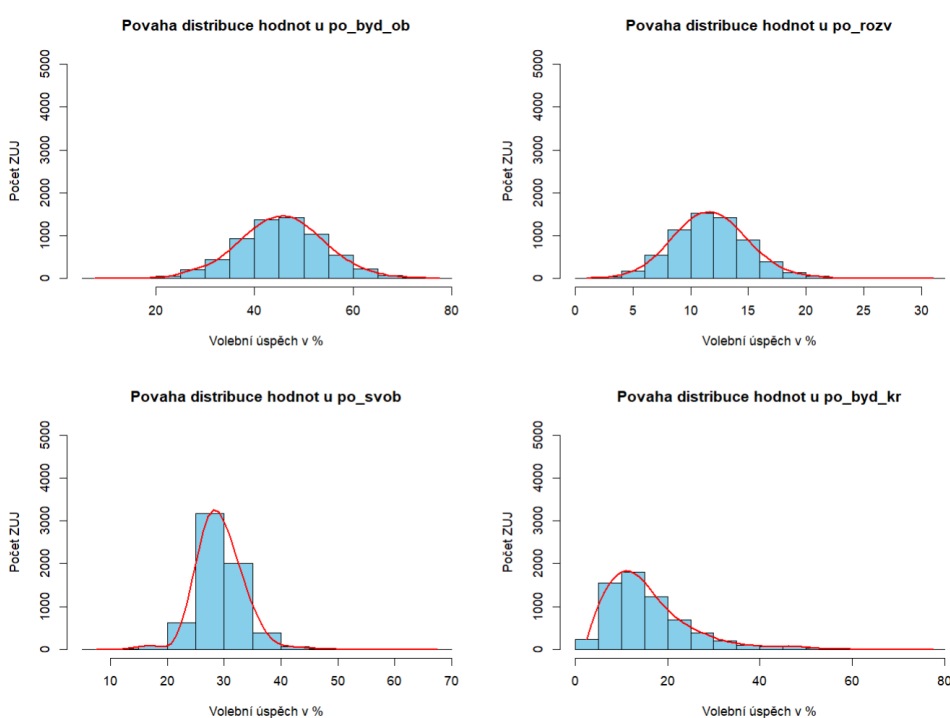
Úspěch hnutí Přísaha na Pohořelicku byl ovlivněn zejména původem zakladatele hnutí Roberta Šlachty z tohoto regionu a neúspěch ČSSD pak přechodem sociálně demokratických voličů k novým politickým subjektům jako ANO 2011 a SPD. Podobné závěry jsou vyvozeny také Lyskem (2021) a jsou obohaceny o další poznatky získané z prezidentských voleb 2023.

### Test normality dat

Dalším krokem po srovnání volebního výsledku politických subjektů bylo provedení testu normality dat pomocí Kolmogorova–Smirnovova testu (dále jen K–S test), který je využíván pro datové sady velkého rozsahu (Lopatecki 2023). Zjištění povahy rozdělení dat je důležité z hlediska využitelnosti určitých statistických metod, které mohou vyžadovat normální rozdělení. Příkladem budiž Pearsonův korelační koeficient, fungující lépe na datech s normálním rozdělením (Schober a kol. 2018).

Z výsledků K-S testu vyplývá, že distribuce většiny atributů nevychází z normálního rozdělení, neboť p-hodnota, kterou je představována hladina významnosti pro zamítnutí nulové hypotézy, nebyla vyšší než 0,05. Výjimkou byly pouze atributy SED ukazatelů podíl rozvedených osob (**po\_roz**) a podíl osob s trvalým bydlištěm v obci trvalého bydliště (**po\_byd\_ob**). Výsledky K-S testů byly zapsány do tabulky.

Za účelem vizuální analýzy normálního rozdělení byly výše zmíněné atributy s normálním rozdělením porovnány s atributy, kde nebyla nulová hypotéza zamítnuta. Pro porovnání byly vybrány atributy podíl svobodných (**po\_svob**) a podíl osob s trvalým bydlištěm mimo kraj obvyklého bydliště (**po\_byd\_kr**). Tyto atributy byly součástí stejné kategorie, jako atributy s diagnostikovaným normálním rozdělením. Při porovnání histogramů byla potvrzena správnost provedeného K-S testu, neboť rozložení příslušných hodnot oproti atributům vybraným pro porovnání odpovídá normální distribuci.



Obr. 12 Porovnání distribuce hodnot u atributů s normálním a odlišným rozdělením

### Korelační analýza

V rámci korelační analýzy byly zkoumány vztahy SED ukazatelů a volebních výsledků. Pro komplexní popis vztahů atributů byly vytvořeny dvě rozsáhlé korelační matice, které byly základem pro výběr vhodných, vzájemně nekorelovaných atributů při zpracování dalších prostorových analýz. Vzhledem k výsledkům provedeného K-S testu byl k výpočtu korelačního koeficientu využit Spearmanův korelační koeficient, který je vhodnější volbou pro data jiné než normální distribuce (Schober a kol. 2018). Tyto korelační matice jsou součástí diplomové práce (Přílohy 1 a 2).

Generalizovanou verzí korelační matice jsou popsány vztahy výše zmíněných politických subjektů s vybranými SED ukazateli. V matici je ukázán kladný vztah mezi podílem římských katolíků v ZUJ a nárůstem voličů Pavla Fishera, nebo záporný vztah mezi podílem vysokoškolsky vzdělaného obyvatelstva a volebním úspěchem hnutí ANO. Zajímavým zjištěním bylo, že v korelační matici nejsou zřetelně zobrazeny vztahy mezi SED ukazateli a hnutím Prísaha. To znamená, že volba Prísahy byla pravděpodobně ovlivněna



odlišnými ukazateli, například osobními preferencemi. Dále byl pozorován opačný typ vztahu některých ukazatelů, například u indexu kvality života byl zjištěn kladný vztah s koalicemi Pirátů a Starostů nebo SPOLU, naopak záporný výsledek byl zpozorován u hnutí ANO 2011 a SPD. Tato skutečnost je důkazem vyhraněných politických preferencí na základě charakteristik regionu pomocí SED. Příkladem budiž region Ústeckého kraje s obecně nižšími hodnotami indexu kvality života a vyšším úspěchem hnutí ANO 2011 a SPD.

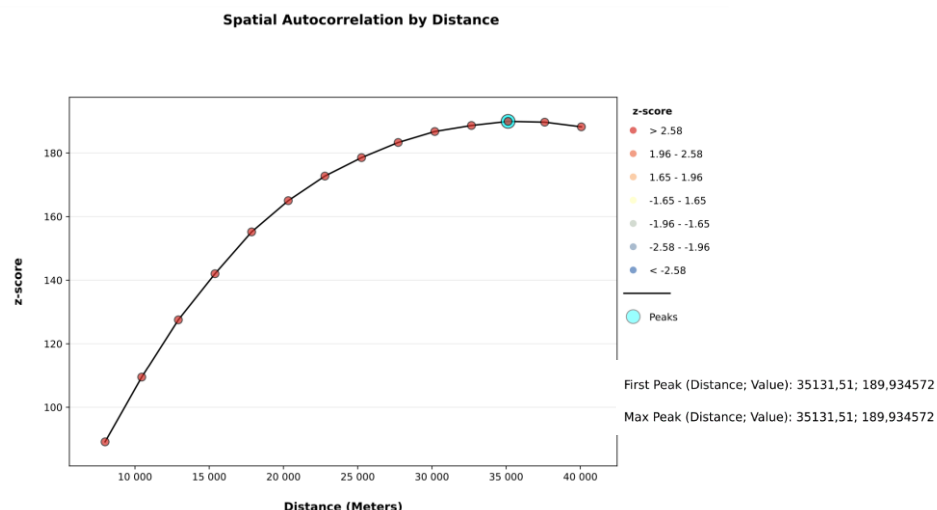


Obr. 13 Spearmanův korelační koeficient vybraných politických subjektů a SED ukazatelů

### Analýza prostorové autokorelace

Závěrečná část exploratorní analýzy byla zaměřena na zjištění míry prostorové autokorelace a identifikace prostorového uspořádání u čtyř politických subjektů s nejvyšší mírou postihnuté variability. Prostorová autokorelace byla zjišťována u koalice SPOLU a hnutí Přísaha pro sněmovní volby. V rámci prezidentských voleb byly vybráni kandidáti Andrej Babiš a Pavel Fisher. Tyto politické subjekty byly vybrány z důvodu postihnutí nejvyšší míry variability volebního chování bez nutnosti provedení analýzy prostorové autokorelace pro všechny kandidáty a politická uskupení.

V prvním kroku analýzy byl využit nástroj *Incremental Spatial Autocorrelation*, jehož výstupem je určení optimální šířky pásma pro zjištění prostorové autokorelace. Optimální šířka pásma je určena na základě prvního vrcholu grafu. Vrcholem grafu je představován průnik z-skóre s definovaným vzdálenostním pásmem v metrech. Z-skóre lze v kontextu grafu chápat, do jaké míry je pozorovaný prostorový vzorec odlišný od náhodného rozložení (Esri 2024a). V rámci lepší porovnatelnosti byla pro všechny atributy využita jednotná šířka vzdálenostního pásma stanovená na 35 km.



Obr. 14 Vrchol grafu, zjištěný nástrojem *Incremental Spatial Autocorrelation* v *ArcGIS Pro*. Následně bylo přistoupeno k výpočtu globálního Moranova I za využití jednotné šířky vzdálenostního pásma a metody euklidovské inverzní vzdálenosti. Touto metodou byl přisouzen vyšší význam geograficky bližším prvkům v datovém souboru vůči prvkům vzdálenějším. Dále byla povolena v rámci řádků standardizace, která je doporučována v případech, kdy je distribuce dat zrcelena v závislosti na prostorovém vzoru (Esri 2024b). Výsledky výpočtu Moranova I byly v další části zaznamenány do tabulky. U všech vybraných politických subjektů byla za pomoci globálního Moranova I diagnostikována tendence k vytváření shluků v rámci prostoru, lišící se pouze určenou mírou, které odpovídá **z-skóre**. U všech subjektů lze zároveň na základě hladiny významnosti (**p-hodnoty**) zamítnout nulovou hypotézu o náhodném uspořádání dat v prostoru.

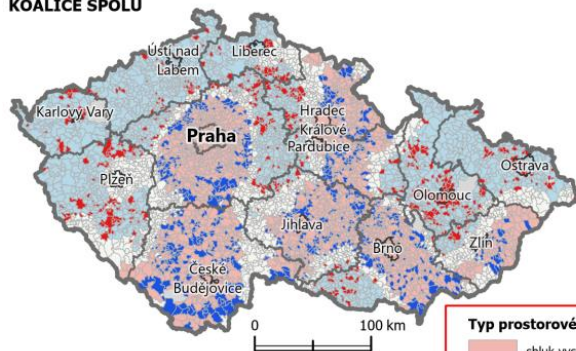
Tab. 2 Výsledky Globálního Moranova I

	Analýza globální prostorové autokorelace pomocí Moranova I		
	Moranův index	z-skóre	p-hodnota
<b>koalice SPOLU</b>	0,2551	186,0000	< 0,00001
<b>hnutí Přísaha</b>	0,1257	91,8245	< 0,00001
<b>Andrej Babiš</b>	0,3005	234,3291	< 0,00001
<b>Pavel Fisher</b>	0,2111	164,7322	< 0,00001

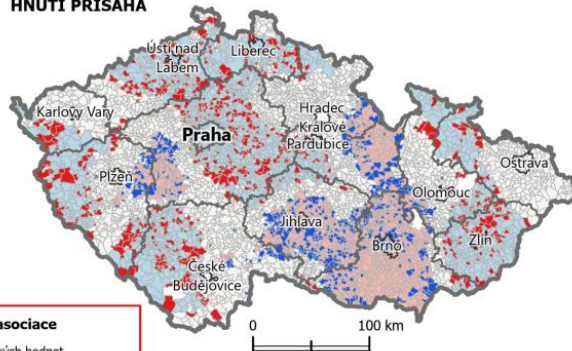
Posledním krokem explorační analýzy bylo zhodnocení lokální prostorové autokorelace. Za účelem diagnózy byla využita lokální verze Moranova I se stejným jednotným nastavením jako v případě výpočtu globálního Moranova I. Z výstupů analýzy byla vytvořena série map, jež posloužila jako příklad pro zpracování budoucích analýz. Nově nalezená zjištění byla dalším potvrzením již dříve zmíněných poznatků, například shluk vysokých hodnot podpory koalice SPOLU v okolí Prahy a Brna.

## ANALÝZA LOKÁLNÍ PROSTOROVÉ AUTOKORELACE POMOCÍ LISA v České republice v letech 2021 a 2023

**KOALICE SPOLU**



**HNUTÍ PŘÍSAHA**

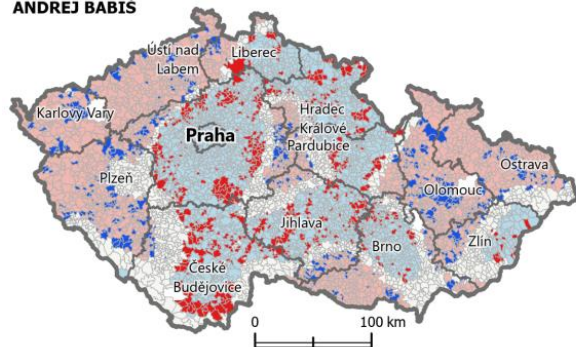


sněmovní volby 2021

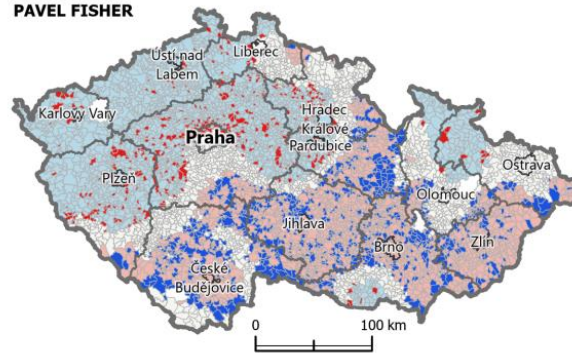
prezidentské volby 2023



**ANDREJ BABIŠ**



**PAVEL FISHER**



Obr. 15 Lokální prostorová autokorelace zjištěná pomocí lokálního Moranova I

## 4.2 Aplikace prostorových analýz

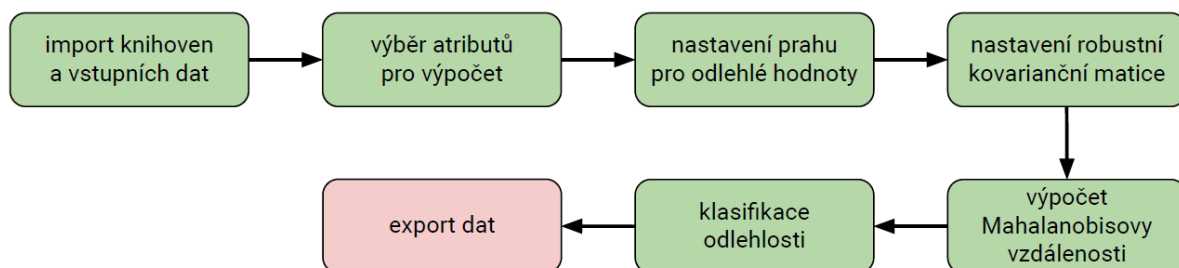
Exploratorní analýzou byla potvrzena tendence k vytváření charakteristických prostorových vzorů v rámci ČR. Prostorové analýzy byly aplikovány za účely mnohorozměrné detekce odlehlosti, tvorby modelů popisujících volební chování nebo vymezení shluků s podporou určitých politických subjektů.

### 4.2.1 Mahalanobisova vzdálenost

První aplikovanou metodou byla detekce odlehklých hodnot v rámci definovaného počtu atributů datové sady. K tomuto účelu byla využita metoda Mahalanobisovy vzdálenosti. Detekce odlehlosti byla aplikována na oba typy voleb nejprve z globálního a poté i lokálního pohledu. Pro obě části analýzy byl vytvořen skript v jazyce R.

#### Globální Mahalanobisova vzdálenost

Nejprve byla zjišťována odlehlost z globálního pohledu, kdy byla vypočítána Mahalanobisova vzdálenost pro všechny záznamy datové sady v definovaném rozsahu atributů. Výpočet byl proveden s využitím robustní verze Mahalanobisovy vzdálenosti, jejíž výhodou je vyšší odolnost vůči odlehlým hodnotám a nevyváženým datům (Li a kol. 2019). Za tímto účelem byla vypočtena robustní kovarianční matice funkcí *covMcd* z balíku *rrcov* (Todorov 2024). Využitím robustní verze Mahalanobisovy vzdálenosti budou odhaleny nejvíce odlehlé záznamy v rámci datového souboru, což je vzhledem k velikosti datového souboru žádoucí.



Obr 16 Řešení výpočtu globální Mahalanobisovy vzdálenosti

Prahová hodnota odlehlosti byla definována jako 95% kvantil Chí kvadrátu. Tato skutečnost znamená, že jako odlehlé hodnoty budou klasifikovány záznamy, jejichž hodnota vypočtené Mahalanobisovy vzdálenosti se nachází nad 95 % kvantilu Chí kvadrát rozdělení. V rámci prahu byly také definovány stupně volnosti s hodnotou 11 pro sněmovní volby a hodnotou 8 pro prezidentské volby. Stupně volnosti jsou parametry Chí kvadrát rozdělení a jsou určovány jako počet vstupních sloupců pro výpočet zmenšený o 1 (Ganti 2024).

Tab. 3 Vstupní atributy pro výpočet globální Mahalanobisovy vzdálenosti

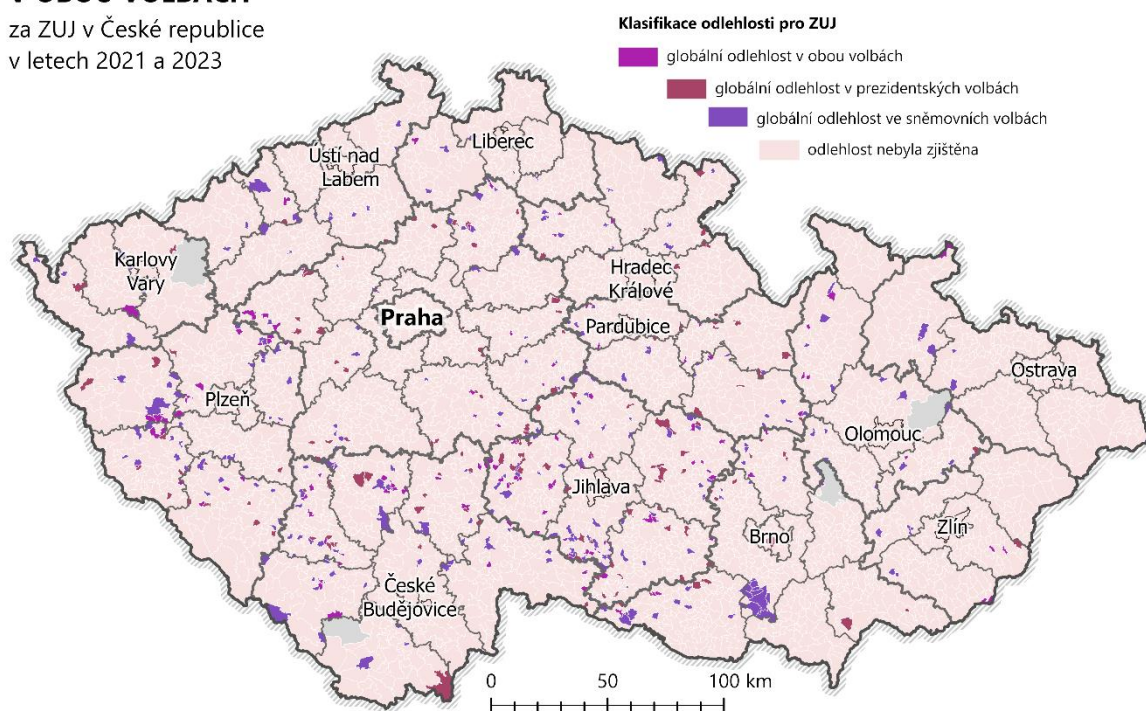
Vstupní atributy pro výpočet globální Mahalanobisovy vzdálenosti	
sněmovní volby 2021	prezidentské volby 2023
SPOLU	Petr Pavel
ANO	Andrej Babiš
Piráti a Starostové	Danuše Nerudová
SPD	Pavel Fisher
Přísaha	Jaroslav Bašta
ČSSD	Marek Hilšer
Volný Blok	Karel Diviš
Zelení	Tomáš Zima
KSČM	volební účast–prezident
TSS	
ostatní politické subjekty	
volební účast–sněmovna	

Definované číslo stupňů volnosti je ovlivněno odlišným počtem vstupních záznamů pro prezidentské a sněmovní volby. Vlivem vysokého počtu vstupních atributů u sněmovních voleb 2021 bylo přistoupeno ke sloučení podílu volebních výsledků u méně úspěšných politických subjektů do atributu **OSTATNI**.

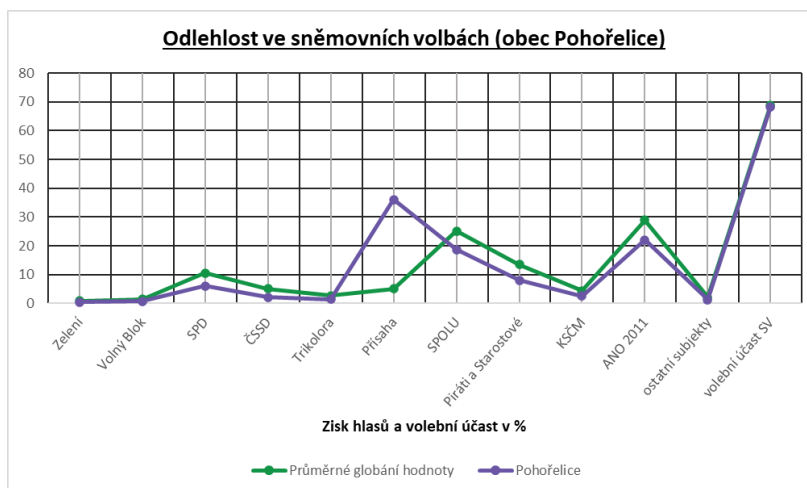
Po výpočtu byla data spojena s prostorovou vrstvou a vytvořena mapa Globální odlehlé hodnoty v obou volbách (Příloha 3). Z detekovaných odlehlých hodnot byly vybrány příklady, na nichž byla odlehlost demonstrována prostřednictvím nejvyšších rozdílů průměrných hodnot a hodnot daného záznamu.

### GLOBÁLNÍ ODLEHLÉ HODNOTY V OBOU VOLBÁCH

za ZUJ v České republice  
v letech 2021 a 2023

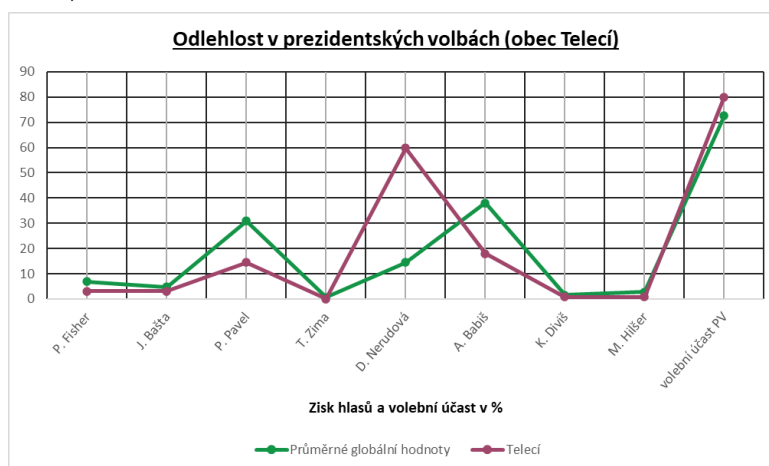


Obr. 17 Mapa globálních odlehlých hodnot na základě Mahalanobisovy vzdálenosti sněmovních i prezidentských voleb

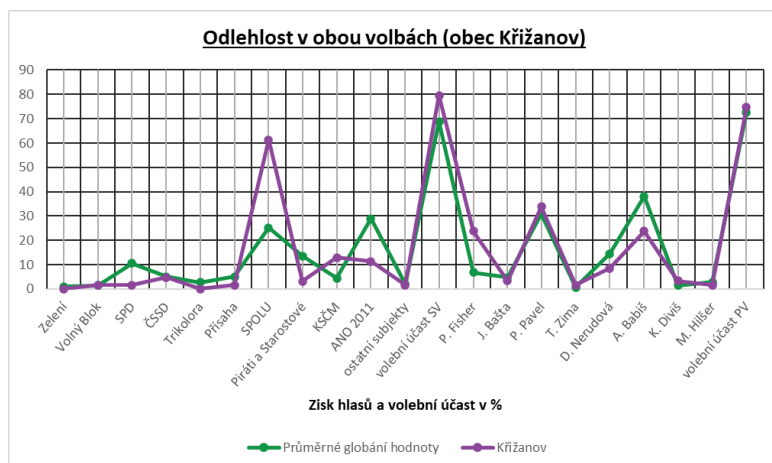


Obr. 18 Porovnání odlehlosti obce Pohořelice a průměrných hodnot ve sněmovních volbách  
Prvním příkladem odlehlosti je obec Pohořelice ve sněmovních volbách a nadprůměrný podíl procentuálních hlasů hnutí Přísaha. Úspěch hnutí je nepochybně ovlivněn faktem původu zakladatele a předsedy hnutí, Roberta Šlachty, z tohoto regionu. Hnutí Přísaha bylo vítězem také v několika přilehlých obcích, čehož je důkazem klasifikace okolních obcí jako odlehlych ve sněmovních volbách 2021.

Druhým příkladem je obec Telecí ležící na severovýchod od Jihlavy s nadprůměrným ziskem volebních hlasů pro Danuši Nerudovou a podprůměrnými výsledky pro Petra Pavla a Andreje Babiše. Úspěch Danuše Nerudové je nepochybně ovlivněn i pozitivní zkušeností místních občanů s kandidátkou v osobní rovině, neboť v této obci je umístěna chalupa rodiny Nerudových, která zde byla před několika lety k rekreačním účelům zakoupena (Nádvořníková 2023).



Obr. 19 Porovnání odlehlosti obce Telecí a průměrných hodnot v prezidentských volbách  
Posledním příkladem je obec Křižanov, která byla klasifikovaná jako odlehlá ve sněmovních i prezidentských volbách. Křižanov se nachází severně od Českých Budějovic v Jihočeském kraji. Důvodem odlehlosti ve sněmovních volbách je nadprůměrný úspěch koalice SPOLU, a naopak neúspěch hnutí ANO 2011. Hlavním důvodem klasifikace ZUJ jako odlehlé v prezidentských volbách byl neúspěch Andreje Babiše. Z tohoto zjištění lze usoudit, že volební chování v Křižanově se mezi volbami nezměnilo, neboť vítězného kandidáta prvního kola, kterým byl Petr Pavel, podpořila v prezidentské kampani také koalice SPOLU.



Obr. 20 Porovnání odlehlosti obce Křižanov a průměrných hodnot v obou volbách

### Lokální Mahalanobisova vzdálenost

Při analýze odlehlosti z lokálního pohledu byla Mahalanobisova vzdálenost vypočtena pouze pro definovaný okruh nejbližších sousedů vybraného prvku v geografickém prostoru, nikoliv pro všechny prvky vstupních dat. Hlavním přínosem metody je detekce odlehlosti pro prvky, jež z globálního pohledu nejsou klasifikovány jako odlehlé, ale při nižším počtu vstupních prvků může být výsledek klasifikace odlišný. V tomto případě byla využita klasická verze Mahalanobisovy vzdálenosti. Důvodem je nižší počet vstupních prvků při výpočtu lokální odlehlosti, z čehož vyplývá potřeba vyšší míry citlivosti ke klasifikaci odlehlých hodnot. K vypočtení lokální Mahalanobisovy vzdálenosti byla využita funkce *Moutlier* z balíku *rrcov*. Výhodou oproti předchozímu řešení je předem definovaný výpočet kovarianční matice v rámci této funkce. Při sestavování skriptu pro výpočet lokální Mahalanobisovy vzdálenosti byly objeveny dva hlavní problémy, které bylo potřeba vyřešit. Prvním problémem byla definice okolí s nedostatečným počtem prvků pro výpočet. V případě nízkého počtu prvků nebyl výpočet proveden a v terminálu bylo zobrazeno chybové hlášení. Tento problém byl vyřešen definicí okolí s dostatečným počtem nejbližších sousedů. Výsledný počet prvků byl po dvaceti testovacích pokusech stanoven na 220 nejbližších sousedů. Alternativou v tomto případě bylo využití fixní vzdálenosti k vyhledání prvků v definovaném okolí. Tato metoda však nebyla při sestavování skriptu vyhodnocena jako vhodná, neboť bylo zapotřebí definovat rozsáhlé okolí v příhraničních ZUJ, a naopak menší okolí u vnitrozemských ZUJ. Následkem definice okolí s příliš vysokým počtem prvků u vnitrozemských ZUJ byla ztráta lokální variability v rámci jednotlivých atributů vstupujících do výpočtu lokální Mahalanobisovy vzdálenosti.

```
>
> # Výpočet Mahalanobisovy vzdálenosti s použitím funkce Moutlier
> moutlier_result <- Moutlier(maha_columns, qchisq = 0.95)
Error in covMcd(X) : n <= p -- you can't be serious!
> |
```

Obr. 21 Chybové hlášení z důvodu nedostatečného počtu prvků

Druhým problémem byl vznik singulární kovarianční matice při výpočtu lokální Mahalanobisovy vzdálenosti. V kovarianční matici jsou zaznamenány vzdálenosti mezi vybraným záznamem a středem datového souboru vybraného atributu. Při výpočtu je pracováno s inverzí této matice, která nemůže být provedena, pokud je alespoň jeden záznam nebo sloupec lineárně závislý na ostatních řádcích nebo sloupcích. Z tohoto důvodu bylo přistoupeno k omezení rozsahu vstupních atributů. V rámci testování možných variant bylo vyzkoušeno také sloučení méně úspěšných

politických subjektů do jednoho atributu, ale po spuštění skriptu bylo v terminálu zobrazeno stejné chybové hlášení o vzniku singulární kovarianční matice.

```

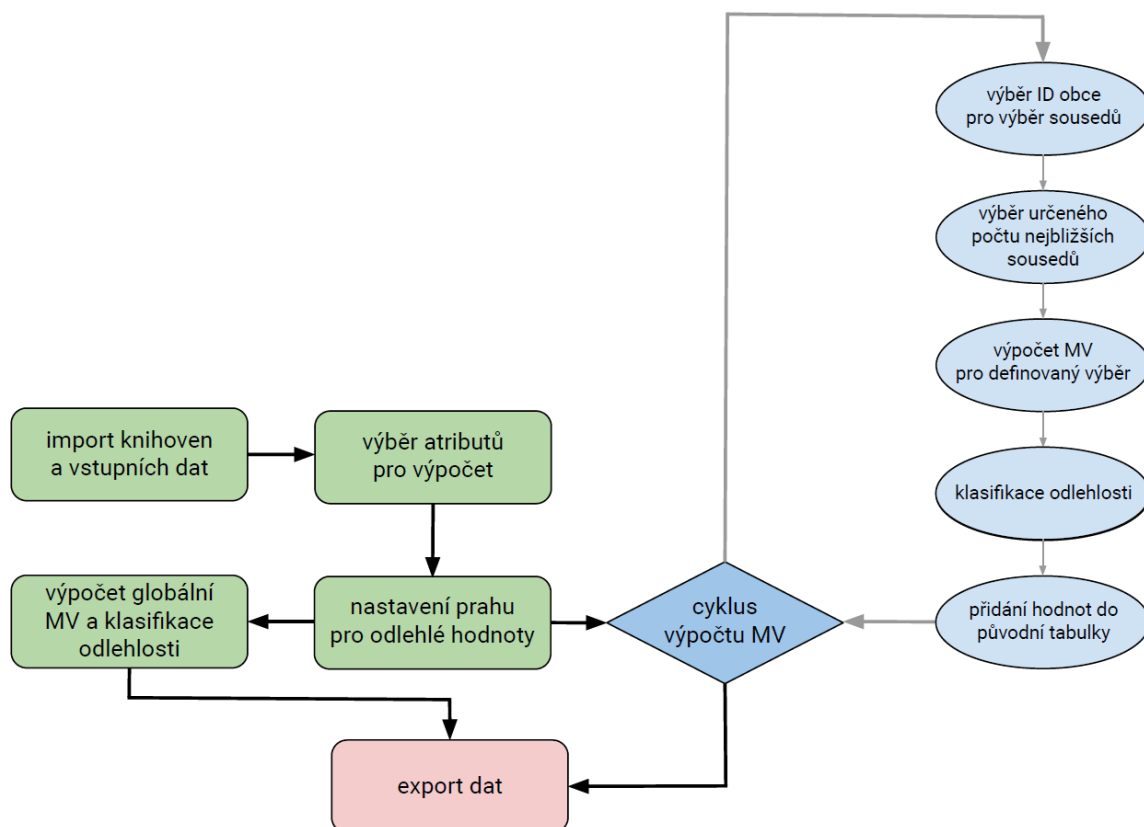
> # Výpočet Mahalanobisovy vzdálenosti s použitím funkce Moutlier
> moutlier_result <- Moutlier(maha_columns, qchisq = 0.95)
Error in solve.default(cov, ...) :
  system is computationally singular: reciprocal condition number = 4.94706e-20
In addition: Warning message:
In covMcd(X) : The covariance matrix of the data is singular.
There are 220 observations (in the entire dataset of 220 obs.) lying on the
hyperplane with equation a_1*(x_i1 - m_1) + ... + a_p*(x_ip - m_p) = 0 with (m_1,
..., m_p) the mean of these observations and coefficients a_i from the vector a <-
c(-0.2338417, -0.4704715, -0.2366298, -0.2366298, -0.2366298, -0.2366298,
-0.2366298, -0.2366298, -0.2366298, -0.2366298, 0.2338417, -0.4704715, 0)

```

Obr. 22 Chybové hlášení z důvodu vzniku singulární kovarianční matice

Po vyřešení výše popsaných problémů bylo přistoupeno k tvorbě konečné podoby skriptu. Před provedením samotného výpočtu lokální Mahalanobisovy vzdálenosti byla vypočtena z důvodu lepší porovnatelnosti s lokální odlehlostí také klasická verze globální Mahalanobisovy vzdálenosti. V následujících krocích již bylo přistoupeno k samotnému výpočtu a detekci lokální odlehlosti prostřednictvím vytvořeného skriptu.

Nejprve byl vytvořen seznam identifikátorů ze sloupce obec\_i\_cast. V dalším kroku byl započat cyklus, jehož úkolem byl nejprve výběr 220 nejbližších sousedů pro daný záznam. Následně byl proveden samotný výpočet lokální Mahalanobisovy vzdálenosti pro vybrané atributy. Nakonec byla klasifikována odlehlost podle stanoveného prahu, jenž byl nastaven stejně jako v případě globální Mahalanobisovy vzdálenosti.



Obr. 23 Řešení výpočtu globální Mahalanobisovy vzdálenosti

V tomto případě bylo z výše popsaných důvodů nastaveno 11 vstupních atributů pro sněmovní volby a 6 pro prezidentské volby. Výpočet Mahalanobisovy vzdálenosti s těmito atributy byl proveden v globální i lokální verzi.



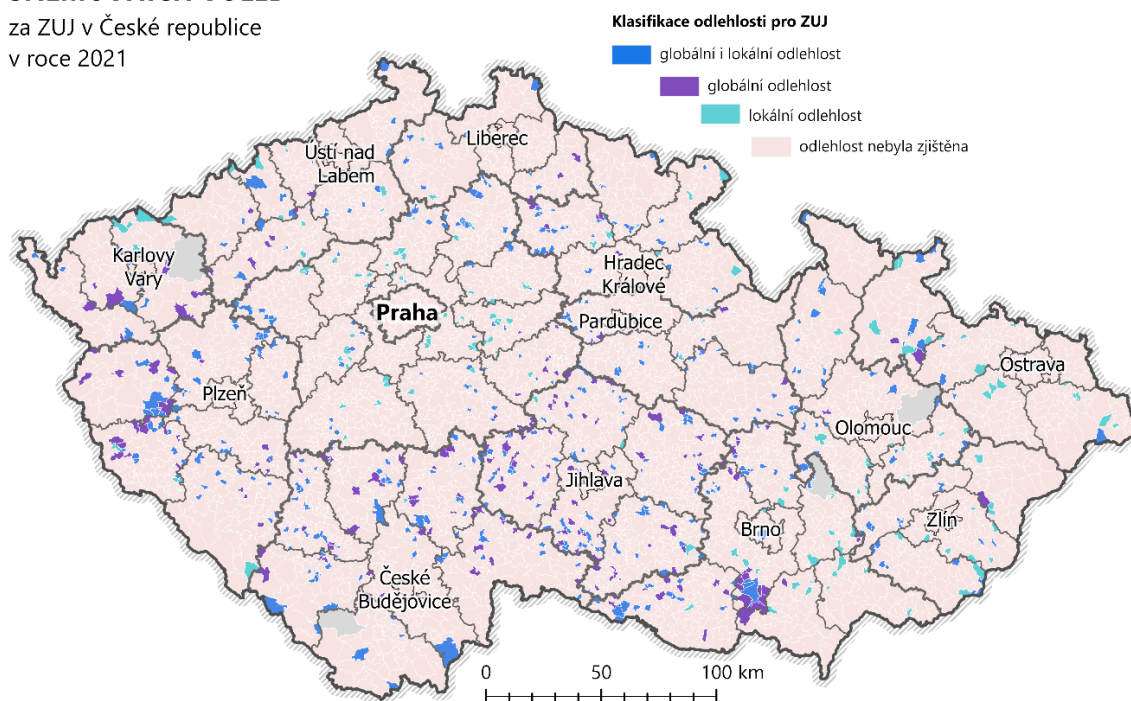
Tab. 4 Vstupní atributy pro výpočet lokální Mahalanobisovy vzdálenosti

Vstupní atributy pro výpočet lokální Mahalanobisovy vzdálenosti	
sněmovní volby 2021	prezidentské volby 2023
SPOLU	Petr Pavel
ANO	Andrej Babiš
Piráti a Starostové	Danuše Nerudová
SPD	Pavel Fisher
Přísaha	Jaroslav Bašta
ČSSD	volební účast–prezident
Volný Blok	
Zelení	
KSČM	
TSS	
volební účast–sněmovna	

Po úspěšném výpočtu lokální Mahalanobisovy vzdálenosti a detekci odlehlosti bylo opět přistoupeno k vizualizaci. V tomto případě byla vytvořena mapa lokálních odlehlých hodnot ve sněmovních volbách (Příloha 4) a mapa lokálních odlehlých hodnot v prezidentských volbách (Příloha 5). Pro porovnání výsledků klasifikace lokální a globální odlehlosti byly vybrány tři modelové příklady.

### ODLEHLÉ HODNOTY SNĚMOVNÍCH VOLEB

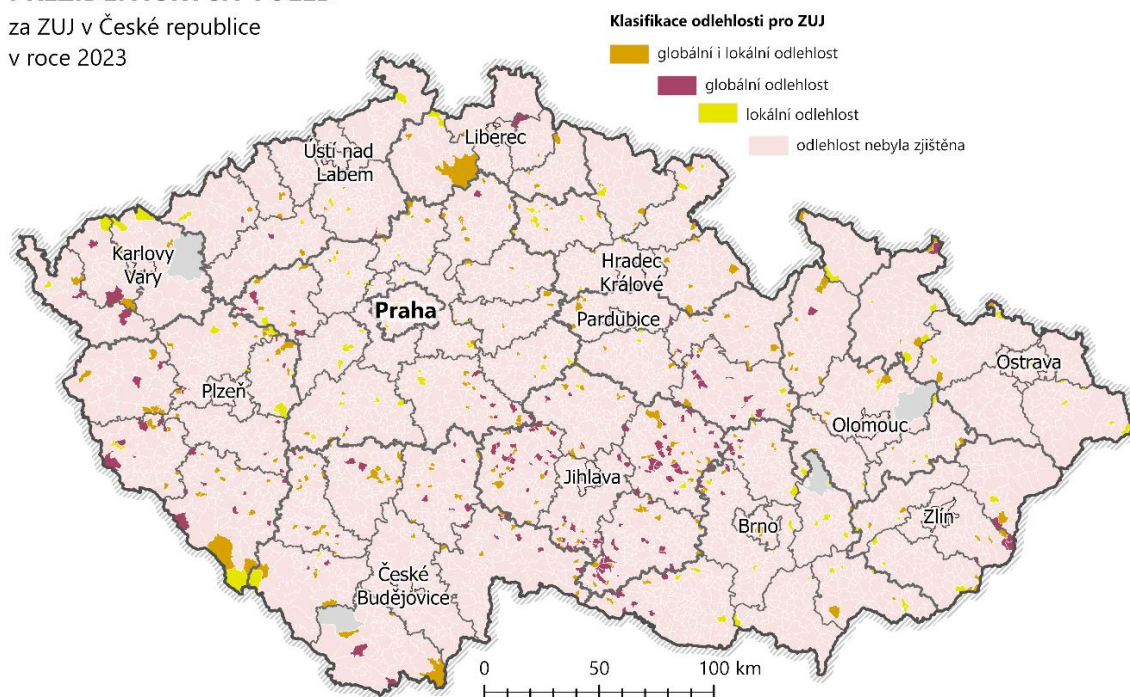
za ZUJ v České republice  
v roce 2021



Obr. 24 Mapa odlehlých hodnot na základě Mahalanobisovy vzdálenosti sněmovních voleb

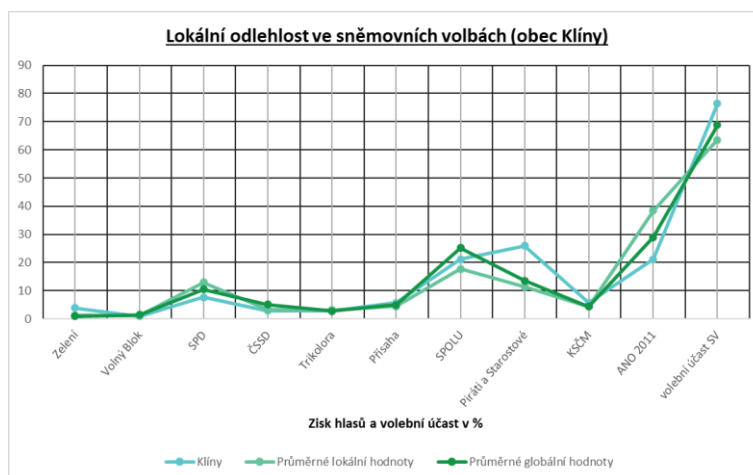
## ODLEHLÉ HODNOTY PREZIDENTSKÝCH VOLEB

za ZUJ v České republice  
v roce 2023



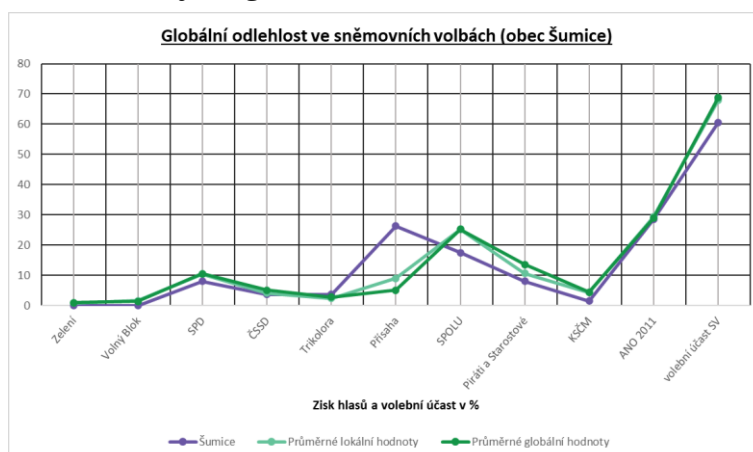
Obr. 25 Mapa odlehlých hodnot na základě Mahalanobisovy vzdálenosti  
prezidentských voleb

Prvním příkladem lokální odlehlosti je obec Klíny, kde hlavním důvodem klasifikace odlehlosti byl nadprůměrný úspěch koalice Pirátů a Starostů a naopak neúspěch hnutí ANO 2011 ve sněmovních volbách. Při porovnání lokálních a globálních průměrných hodnot je viditelný nižší rozdíl mezi hodnotami obce Klíny a průměrnými globálními hodnotami. Klasifikace lokální odlehlosti je v tomto případě způsobena umístěním obce v Ústeckém kraji, který lze z hlediska volebních výsledků označit jako kraj s dominantním postavením hnutí ANO 2011.



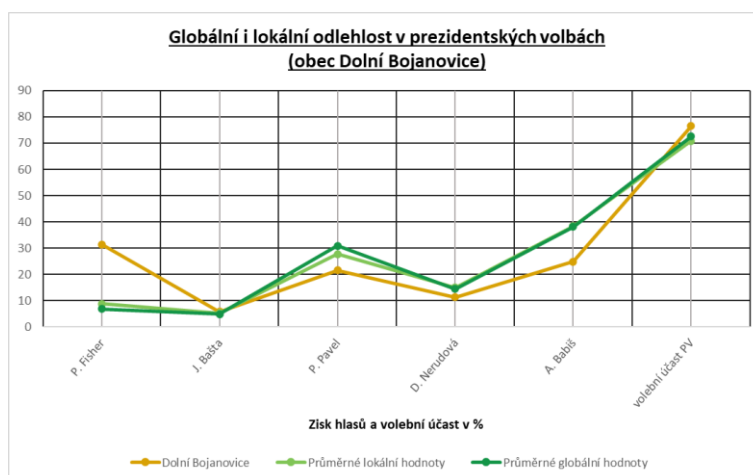
Obr. 26 Porovnání lokální odlehlosti obce Klíny a průměrných lokálních i globálních hodnot  
Jako příklad globální odlehlosti byla vybrána obec Šumice nedaleko Pohorelic. Jak již bylo zmíněno u příkladů s globální Mahalanobisovou vzdáleností, typickým znakem tohoto regionu je úspěch hnutí Přísaha v Pohorelicích a přilehlých obcích. V kontextu rozdílu mezi lokální a globální odlehlostí je viditelný nižší rozdíl mezi průměrnými lokálními hodnotami a hodnotami obce Šumice například u procentuálních podílů volebních výsledků hnutí

Přísaha nebo koalice Pirátů a Starostů. Zajímavostí je klasifikace některých okolních ZUJ, včetně samotných Pohořelic, jako globálně i lokálně odlehlé.



Obr. 27 Porovnání globální odlehlosti obce Šumice a průměrných lokálních i globálních hodnot

Jako poslední příklad byla vybrána obec Dolní Bojanovice za prezidentské volby 2023, která je v kontextu volebního chování i nadprůměrného zastoupení věřících přezdívána jako moravský Vatikán (Biben 2023). V obci byl vítězem senátor za KDU–ČSL Pavel Fisher s nejvyšším počtem hlasů. Vzhledem k této skutečnosti jsou Dolní Bojanovice jednou z mála obcí, kde nebyl vítězem ani jeden z kandidátů postoupivších do druhého kola prezidentské volby. V kontextu globální a lokální odlehlosti byla obec vyhodnocena jako globálně i lokálně odlehlá, což dokazuje téměř shodné postavení globálních i lokálních průměrných hodnot vůči hodnotám obce.



Obr. 28 Porovnání globální i lokální odlehlosti obce Dolní Bojanovice a průměrných lokálních i globálních hodnot

## 4.2.2 Analýza využití voličských průkazů

Po analýze lokální a globální odlehlosti bylo přistoupeno k analýze využití voličských průkazů. Cílem analýzy bylo vyhledání obcí s vysokou mírou využití voličských průkazů pro lednové prezidentské volby roku 2023. Lednové období konání voleb mělo za následek vyšší míru využití voličských průkazů v oblastech lyžařských středisek.

Nejprve byl podniknut pokus o vyhledání volně dostupných dat evidence voličských průkazů či jiných statistických ukazatelů k tomuto tématu, který však skončil neúspěchem. Proto bylo přistoupeno k diskusi s konzultantem diplomové práce v oblasti politologie. Výsledkem konzultace bylo potvrzení absence datové evidence voličských průkazů, například s prostorovou podrobností ZUJ. Z tohoto důvodu byl navržen alternativní postup pro zjištění míry využití voličských průkazů spočívající v porovnání evidence voličů v seznamu prezidentských voleb 2023 s evidencí voličů v seznamu krajských voleb 2020. Krajské volby 2020 byly zvoleny z důvodu říjnového data konání a regionálního charakteru těchto voleb. Dále bylo doporučeno zohlednění počtu obyvatel obcí za roky 2023 a 2020 pro lepší porovnatelnost s nárůstem či poklesem voličů v seznamu ve sledovaných letech.

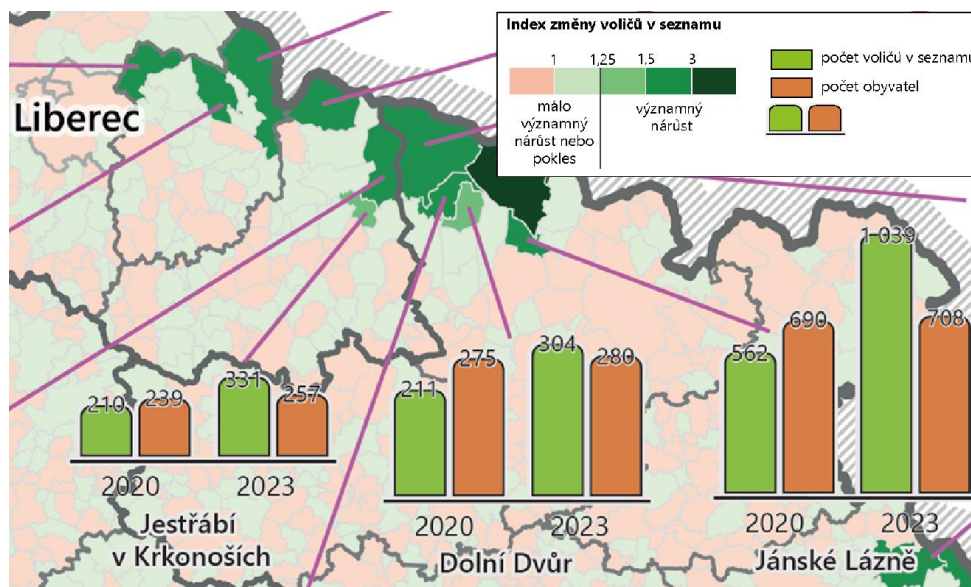
Prvním krokem k provedení analýzy bylo shromáždění potřebných dat voličů v seznamu a počtu obyvatel. Dalším krokem bylo vypočtení indexu změny voličů v seznamu jako podílu hodnot voličů v seznamu za rok 2023 vůči hodnotám voličů v seznamu v roce 2020.

ID_ZUJ	nazev	vol_sez_2023	vol_sez_2020	poc_ob_2023	poc_ob_2020	index_zmeny_vol_sez_2023_2020
554979	Abertamy	703	720	896	939	0,976388889
531367	Adamov	109	110	133	125	0,990909091
535826	Adamov	719	725	1014	963	0,991724138
581291	Adamov	3512	3556	4538	4507	0,987626547
547786	Adršpach	405	409	495	503	0,990220049
547981	Albrechtice	342	355	424	455	0,963380282
598925	Albrechtice	3107	3160	3826	3842	0,983227848
576077	Albrechtice nad Orli	778	797	994	978	0,976160602
549258	Albrechtice nad Vlt.	793	785	1003	918	1,010191083
563528	Albrechtice v Jizers	378	288	394	349	1,3125
568741	Albrechtický	564	578	724	705	0,975778547
506761	Alojzov	206	204	257	259	1,009803922
538001	Andělská Hora	301	301	406	370	1
551929	Andělská Hora	303	320	397	371	0,946875
573426	Anenská Studánka	156	151	187	173	1,033112583
586030	Archebov	694	710	872	882	0,977464789
509388	Arnešovice	64	65	85	85	0,984615385
586854	Arnolec	135	134	192	174	1,007462687
562343	Arnoltice	352	354	422	407	0,994350282
554499	Aš	9443	9497	12804	13182	0,994313994
537241	Babice	103	103	153	133	1
538043	Babice	933	909	1414	1265	1,02640264
552356	Babice	370	371	448	451	0,997304582
569828	Babice	166	166	212	185	1
590274	Babice	174	170	205	194	1,023529412
592013	Babice	1429	1453	1879	1825	0,98348245

Obr.29 Pohled na připravený datový soubor k analýze využití voličských průkazů

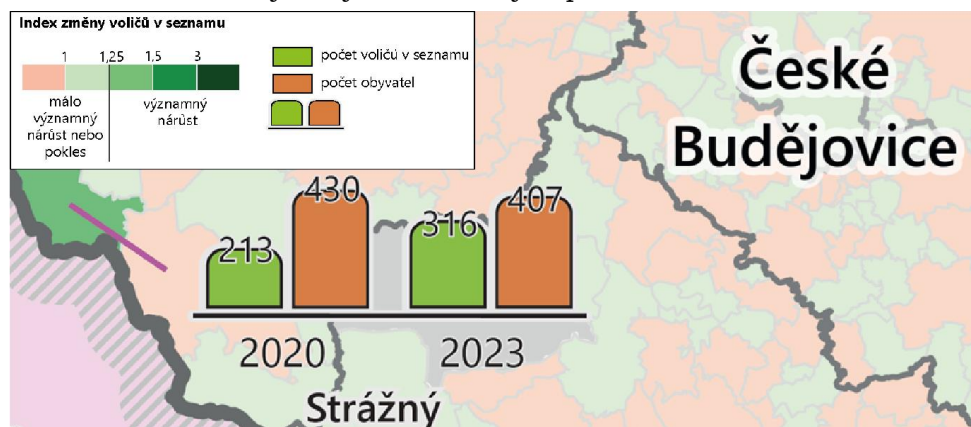
Druhým krokem byla vizualizace shromážděných a vypočtených dat v ArcGIS Pro. Při vizualizaci byla s konzultantem diplomové práce diskutována hranice signifikantní změny u vypočteného indexu. Hranicí signifikantní změny byla nakonec určena hodnota indexu 1,25. Tato hodnota je interpretována jako 25% nárůst voličů v seznamu v roce 2023 oproti roku 2020. Indexem jsou zohledněny také ZUJ s poklesem voličů v seznamu u prezidentských voleb v porovnání s krajskými volbami. Hodnota indexu v tomto případě nabývá hodnot menších než 1. Výsledkem analýzy voličských průkazů je mapa Porovnání voličstva mezi volbami (Příloha 6).

Z vytvořené mapy byla zjištěna signifikantní pozitivní změna vypočteného indexu u 29 ZUJ, odpovídající v tomto případě samostatným obcím v příhraničních horských oblastech s vyhlášenými lyžařskými středisky. K obcím se zjištěnou pozitivní signifikantní změnou indexu byly v dalším kroku přidány sloupcové kartodiagramy zobrazující počet voličů v seznamu a počet obyvatel za odpovídající časové vymezení. Charakteristickým znakem u 29 ZUJ se zjištěnou signifikantní mírou změny vypočteného indexu byly vyšší hodnoty voličů v seznamu za rok 2023, v porovnání s počtem voličů v seznamu 2020 nebo s počty obyvatel za oba zmíněné roky.



Obr. 30 Přehled změn voličů v seznamu a počtu obyvatel mezi lety 2020 a 2023 v Krkonoších

Jedinou ZUJ s diskutabilním zařazením byla obec Strážný v Jihočeském kraji, přestože i v tomto případě byla zjištěna signifikantní míra změny. Důvodem k diskusi o správnosti zařazení je nižší počet voličů v seznamu vůči počtu obyvatel obce pro oba zobrazované roky. Z tohoto důvodu mohou existovat další vlivy, jež způsobily změnu voličů v seznamu za sledované období, jimiž může být náhlá změna věkové struktury obyvatel, vysoký počet přistěhovalých apod. Vzhledem k těmto vlivům není možné zcela potvrdit jako důvod nárůstu voličů v seznamu vyšší využití voličských průkazů.



Obr. 31 Obec Strážný s diskutabilní mírou využití voličských průkazů

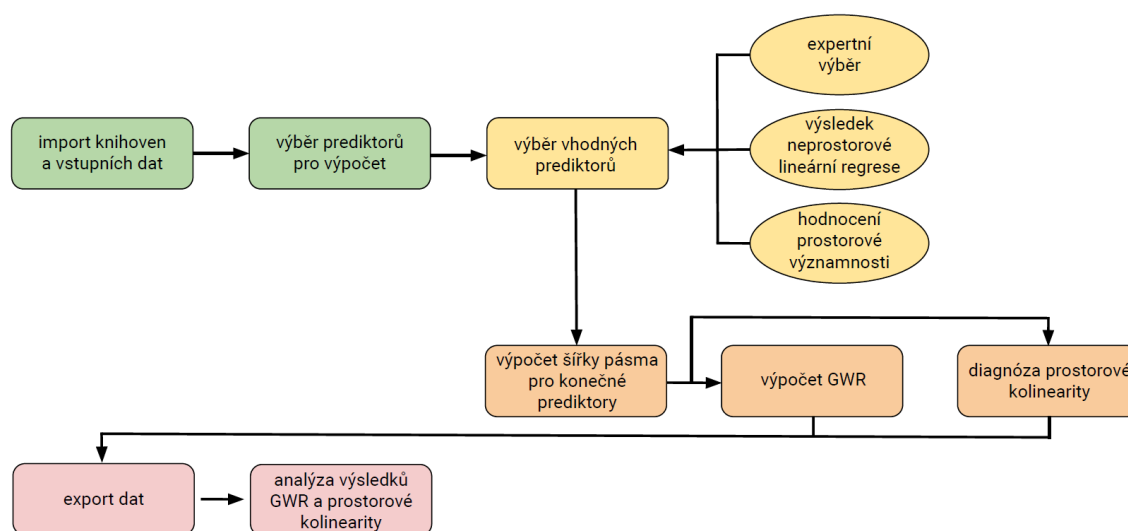
Výsledkem analýzy je nepřímé hodnocení míry využití voličských průkazů prostřednictvím dat o voličích v seznamu za vymezené roky. Ačkoliv se přímo nejedná o práci s daty evidence voličských průkazů, aplikovanými metodami byly identifikovány ZUJ, u kterých byla vyšší míra využití voličských průkazů předpokládána.

### 4.2.3 Geograficky vážená regrese

Ke zpracování třetí analýzy byla využita metoda geograficky vážené regrese. Cílem analýzy byl popis chování čtyř nejúspěšnějších politických subjektů ve sněmovních volbách 2021 jako závislých proměnných s pomocí SED ukazatelů jako jejich prediktorů prostřednictvím sestavených GWR modelů. Hlavní výhodou oproti klasické lineární regresi je odstranění prostorové závislosti, která je v lineárně regresním modelu pozorována na reziduích modelu a prostorová proměnlivost prediktorů vysvětlujících závislou proměnnou. Zpracované regresní modely byly vytvořeny podle metodiky Golliniové a kol. (2015) pracující v prostředí R s balíkem *GWmodel*. Zjevné výhody využití tohoto balíku pro zpracování GWR byly nastíněny v rešeršní části práce.

#### Sestavení skriptu pro výpočet GWR

Prvním krokem pro vytvoření GWR modelů bylo sestavení komplexního řešení, zahrnujícího kromě samotného sestaveného GWR modelu také hodnocení vhodnosti prediktorů neboli vysvětlujících proměnných. Dále je v rámci skriptu vypočtena optimální šířka pásma pro GWR a hodnocení míry lokální kolinearit.



Obr. 32 Vytvoření optimálního GWR modelu s hodnocením jeho lokální kolinearit

K výběru optimálních vstupních prediktorů pro vysvětlení závisle proměnné bylo využito tří hlavních metod. První metodou byl expertní výběr prediktorů na základě předpokladu nejlepšího vysvětlení závislé proměnné. K expertnímu výběru byly dále využity informace získané z exploratorní analýzy, např. vzájemné vztahy SED ukazatelů a dané závislé proměnné zobrazené v korelační matici (Obr. 8). Dále byly vyřazeny prediktory s jasným vzájemným vztahem. Příkladem tohoto vyřazení je využití pouze jednoho atributu popisujícího podíl žen z hlediska počtu narozených dětí, ačkoliv jsou v sestavené datové sadě přítomny čtyři atributy s tímto tématem (např. podíl žen s jedním narozeným dítětem a podíl žen se dvěma narozenými dětmi).

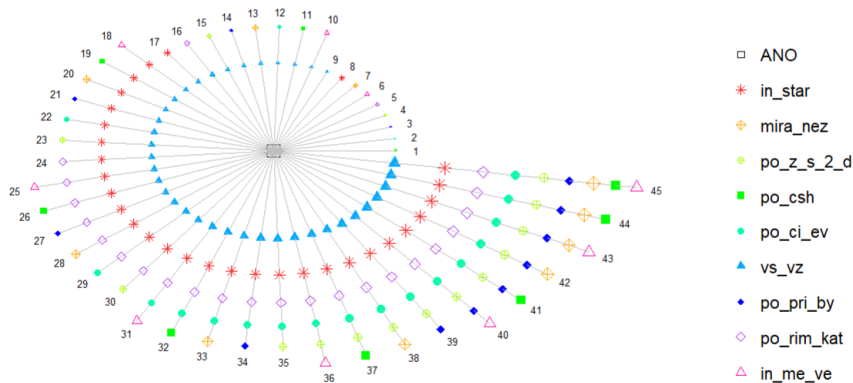
Další využitou metodou byl výsledek klasické neprostorové lineární regrese, kde je zobrazeno hodnocení prediktorů a samotného modelu využitých k výpočtu na základě specifických ukazatelů. Prvním ukazatelem je standardní chyba (**Std. Error**), kterou je vyjádřena míra variability reziduí v okolí regresní přímky, kdy cílem bylo tuto chybu minimalizovat. Druhým ukazatelem je poměr odhadu koeficientu k jeho standardní chybě (**t value**). V tomto případě platí, že čím více je prediktor vzdálený od 0 v kladném či záporném směru, tím vyšší je vliv prediktoru na výsledný model. Třetím ukazatelem je hladina významnosti (**Pr > |t|**) – pokud je hladina významnosti prediktoru nižší než

standardní hladina významnosti pro zamítnutí nulové hypotézy, je prediktor považován za významný v kontextu lineárního regresního modelu. Posledním ukazatelem z hlediska definovaných prediktorů jsou např. **Estimate** popisující hodnotu regresního koeficientu příslušného prediktoru vůči definované závislé proměnné. Definice přesnosti modelu pak vychází z vypočtených maxim a minim reziduí a jejich standardní chyby. Nejdůležitějším ukazatelem je v tomto případě **Multiple R-squared** popisující kvalitu vytvořeného modelu v kontextu výpovědní hodnoty o chování závislé proměnné. Tyto ukazatele jsou také součástí později vytvořeného GWR modelu, ovšem s rozdílným pojmenováním jednotlivých ukazatelů.

Call: lm(formula = formula, data = data_sf)	<b>SPOLU</b>	Call: lm(formula = formula, data = data_sf)	<b>ANO 2011</b>
Residuals: Min IQ Median 3Q Max -32.611 -4.250 -0.268 3.907 39.089		Residuals: Min IQ Median 3Q Max -27.064 -3.883 -0.227 3.683 34.304	
Coefficients: Estimate Std. Error t value Pr(> t ) (Intercept) 20.821152 0.663553 31.378 < 2e-16 *** in_star -0.006049 0.001065 -5.681 1.4e-08 *** mira_nez -0.376864 0.032473 -11.606 < 2e-16 *** sek_s -0.103927 0.010893 -9.540 < 2e-16 *** po_rim_kat 0.211263 0.008096 26.093 < 2e-16 *** vs_vz 0.538979 0.014925 36.114 < 2e-16 ***		Coefficients: Estimate Std. Error t value Pr(> t ) (Intercept) 29.957113 0.594126 50.422 < 2e-16 *** in_star 0.008301 0.001058 7.844 5.09e-15 *** mira_nez 0.217974 0.036193 6.023 1.81e-09 *** po_z_s_2_d 0.066471 0.009356 7.105 1.34e-12 *** po_csh -0.878452 0.178115 -4.932 8.35e-07 *** po_ci_ev -0.327448 0.056763 -5.769 8.36e-09 *** vs_vz -0.471251 0.013007 -36.230 < 2e-16 *** po_pri_by 1.184896 0.109784 10.793 < 2e-16 ***	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1		Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
Residual standard error: 6.62 on 6382 degrees of freedom Multiple R-squared: 0.3787, Adjusted R-squared: 0.3782 F-statistic: 777.9 on 5 and 6382 DF, p-value: < 2.2e-16		Residual standard error: 6.644 on 6380 degrees of freedom Multiple R-squared: 0.2564, Adjusted R-squared: 0.2556 F-statistic: 314.3 on 7 and 6380 DF, p-value: < 2.2e-16	
Call: lm(formula = formula, data = data_sf)	<b>Piráti a Starostové</b>	Call: lm(formula = formula, data = data_sf)	<b>SPD</b>
Residuals: Min IQ Median 3Q Max -17.360 -3.281 -0.658 2.641 40.239		Residuals: Min IQ Median 3Q Max -16.245 -2.501 -0.385 1.966 32.195	
Coefficients: Estimate Std. Error t value Pr(> t ) (Intercept) 15.10304 1.04141 14.503 < 2e-16 *** in_star -0.11372 0.02325 -4.892 1.02e-06 *** po_pri_by -0.49693 0.07414 -6.703 2.22e-11 *** prim_s -0.04657 0.01022 -4.558 5.27e-06 *** po_csh 0.77571 0.13489 5.751 9.31e-09 *** po_ci_ev 0.23146 0.04305 5.377 7.84e-08 *** vs_vz 0.25110 0.01004 25.003 < 2e-16 ***		Coefficients: Estimate Std. Error t value Pr(> t ) (Intercept) 2.945609 0.367473 8.016 1.29e-15 *** mira_nez 0.309819 0.020072 15.436 < 2e-16 *** sek_s 0.096635 0.006358 15.198 < 2e-16 *** po_rozv 0.249149 0.016778 14.850 < 2e-16 *** po_byd_kr -0.076316 0.006360 -12.000 < 2e-16 *** po_z_s_3 0.026595 0.005254 5.062 4.27e-07 ***	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1		Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1	
Residual standard error: 5.038 on 6381 degrees of freedom Multiple R-squared: 0.1535, Adjusted R-squared: 0.1527 F-statistic: 192.9 on 6 and 6381 DF, p-value: < 2.2e-16		Residual standard error: 4.132 on 6382 degrees of freedom Multiple R-squared: 0.1556, Adjusted R-squared: 0.1549 F-statistic: 235.2 on 5 and 6382 DF, p-value: < 2.2e-16	

Obr. 33 Výsledek neprostorových regresních modelů lineární regrese

Třetí metodou, jenž byla v kontextu výběru optimálních prediktorů využita, byla funkce *gwr.model.selection* z balíku *GWmodel*. V rámci této funkce jsou vstupní prediktory seřazeny z hlediska prostorové významnosti. Celý proces je založen na iterativním průchodu a postupném seřazení vstupních prediktorů z hlediska prostorové významnosti.



Obr. 34 Seřazení vstupních prediktorů podle prostorové významnosti nástrojem *gwr.model.selection*

Po určení optimálních prediktorů byla vypočítána optimální šířka vzdálenostního pásma výpočtu GWR modelu pomocí funkce *bw.gwr*. V tomto případě byla na rozdíl od analýzy lokální Mahalanobisovy vzdálenosti zvolena metoda výběru na základě vzdálenosti, nikoliv počtu nejbližších sousedů. Hlavním důvodem bylo zamezení přílišnému přizpůsobení tvaru jádra zejména v okrajových regionech ČR a vzniku zkreslených výsledků GWR v těchto oblastech. Výpočet optimálního jádra GWR modelu je určován pomocí hodnot křížové validace, jež jsou v rámci výpočtu náhodně vybírány. Tento proces je opakován, dokud není dosaženo optimální hodnoty pro šířku vzdálenostního pásma.

Tab. 5 Optimální šířka pásma podle závisle proměnné

Šířka vypočtených pásem	
Závislá proměnná	šířka pásma [m]
SPOLU	48212,20
ANO 2011	48600,00
PIRSTAN	33953,70
SPD	41464,48

Kromě hodnocení vhodnosti vybraných prediktorů byla ze stejného balíku do skriptu přidána také funkce výpočtu lokální kolinearita výsledného modelu. Lokální kolinearita je v kontextu GWR chápána jako míra vzájemné závislosti vstupních prediktorů (Gollini a kol. 2015).

Funkcí *gwr.basic* je vypočítán model GWR s možností nastavení vstupních parametrů. Jedná se například o šířku vzdálenostního pásma nebo volbu jádra pro výpočet vážených vzdáleností. K provedení výpočtu byla využita šířka vzdálenostního pásma z výše popsaného nástroje *gwr.basic* a byla nastavena kvadratická funkce k výpočtu vážené vzdálenosti.

V poslední části skriptu byla vypočtena lokální kolinearita vstupních prediktorů pro výpočet GWR s využitím nástroje *gwr.collin.diagno*. Funkcí je vypočtena míra lokální kolinearit GWR modelu i míra kolinearit mezi použitými prediktory. Vysokou mírou lokální kolinearit je způsobena např. vysoká nestabilita v odhadech prediktorů GWR modelu. Z tohoto důvodu je doporučováno při míře lokální kolinearit vyšší než 30 podniknutí kroků k jejímu snížení, např. omezení počtu vstupních prediktorů (Gollini a kol. 2015), k čemuž bylo při nepříznivých výsledcích opravdu přistoupeno.

Po vytvoření optimálního GWR modelu s příznivou mírou kolinearit byly výsledné modely exportovány ve formátu GPKG. Důvodem bylo příliš dlouhé pojmenování atributů pro formát SHP. Po úpravě pojmenování atributů v programu QGIS byla data exportována do formátu SHP a dále zpracována do podoby map v prostředí ArcGIS Pro.

### Vytvoření GWR modelů

Po sestavení skriptu byly vytvořeny GWR modely vítězů voleb do poslanecké sněmovny 2021, kterými byly koalice SPOLU, hnutí ANO 2011, koalice Pirátů a Starostů a hnutí SPD. Optimální vstupní prediktory jednotlivých politických subjektů byly vybrány s využitím popsaných při tvorbě skriptu GWR.

Po provedeném testování byly nakonec vybrány prediktory s nejlepšími výsledky z hlediska expertního výběru, lineární regrese, prostorové významnosti a lokální kolinearit. V rámci komplexního procesu výběru byly postupně vyřazovány prediktory s nízkou mírou vysvětlovací hodnoty pro regresní model. Při výpočtu GWR modelů byla dále zjištěna rozdílná lokální kolinearita a míra vysvětlení variability ukazatelem  $R^2$  podle počtu vstupních prediktorů. Z tohoto důvodu byly nakonec vytvořeny GWR modely s rozdílným počtem vstupních prediktorů. Tento jev byl způsoben vzájemnými vztahy prediktorů, které byly zjištěny při exploratorní analýze z korelační matice SED ukazatelů (Příloha 2).

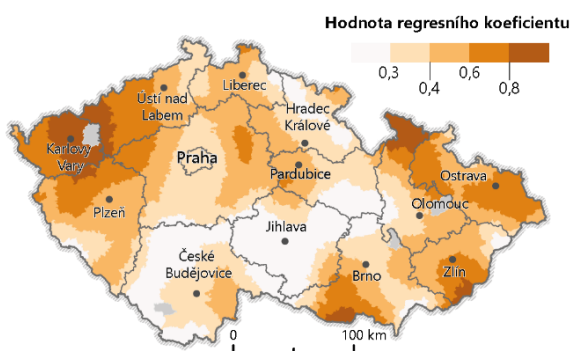


Tab. 6 Vybrané optimální prediktory k vytvoření GWR modelu podle závislých proměnných

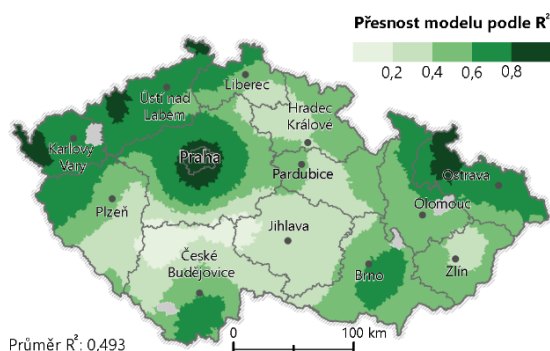
Vstupní atributy pro výpočet GWR modelů			
SPOLU	ANO	Piráti a Starostové	SPD
index stáří	index stáří	index stáří	míra nezaměstnanosti
míra nezaměstnanosti	míra nezaměstnanosti	podíl primárního sektoru	podíl sekundárního sektoru
podíl sekundárního sektoru	podíl žen se dvěma dětmi	podíl věřících husitské církve	podíl rozvedených osob
podíl věřících římskokatolické církve	podíl věřících husitské církve	podíl věřících evangelické církve	podíl osob s OB mimo kraj TB
podíl vysokoškolsky vzdělaných	podíl věřících evangelické církve	podíl vysokoškolsky vzdělaných	podíl žen se třemi a více dětmi
	podíl vysokoškolsky vzdělaných	podíl obyvatel s příspěvkem na bydlení	
	podíl obyvatel s příspěvkem na bydlení		

Prvním výstupem byl Prostorový regresní model koalice SPOLU (Příloha 7). GWR model byl vypočten z pěti vstupních prediktů, kterým bylo kupříkladu podíl vysokoškolsky vzdělaných. V prostoru je viditelný nejvýznamnější pozitivní vztah v okolí Karlových Varů nebo na Jesenicku. Hodnota regresního koeficientu v tomto případě znamená, že se zvýšením hodnoty podílu vysokoškolsky vzdělaných v těchto oblastech bude zaznamenán nárůst také ve volebním výsledku koalice SPOLU. Vytvořeným modelem je nejlépe popisován výsledek koalice SPOLU v okolí Prahy, kde byl zaznamenán úspěch koalice, a také v příhraničních regionech ČR, kde byl volební výsledek koalice slabší. Na pozorovaném území je vyšší lokální kolinearitou než 30 ovlivněno 21,95 % všech územních prvků. Ačkoliv se jedná o GWR model s nejvyšší mírou kolinearit, z hlediska přesnosti modelu je modelem vhodně vysvětleno prostorové chování závislé proměnné.

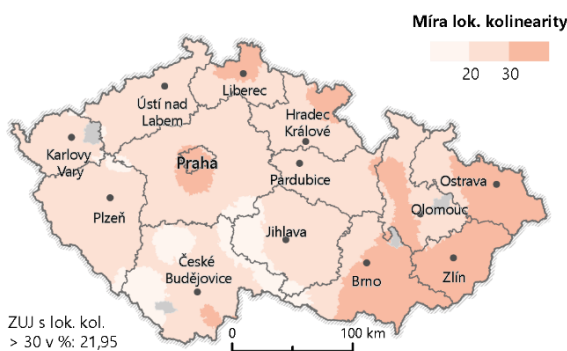
#### PODÍL VYSOKOŠKOLSKY VZDĚLANÝCH



#### PŘESNOST REGRESNÍHO MODELU



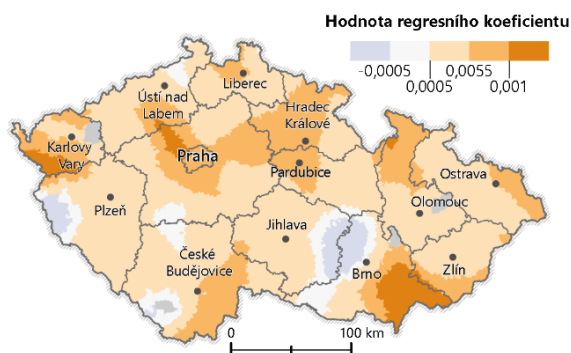
#### PROSTOROVÁ KOLINEARITA



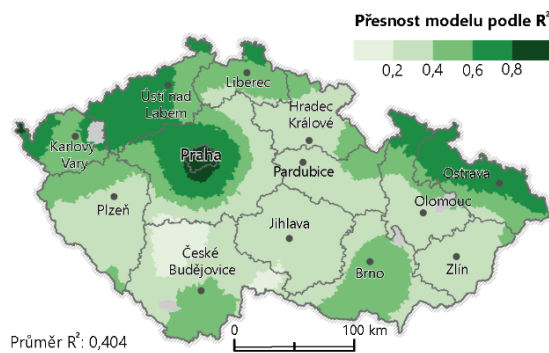
Obr. 35 Vybrané ukazatele GWR modelu pro koalici SPOLU

Druhý GWR model byl vypočten s využitím sedmi vstupních prediktorů pro hnutí ANO 2011, které se ve sněmovních volbách 2021 umístilo na druhém místě (Příloha 8). Počet vstupních regresních koeficientů je v tomto případě nejvyšší ze všech GWR modelů. Důvodem je nižší míra lokální kolinearit v porovnání s GWR modelem koalice SPOLU a porovnatelná míra vysvětlení variability modelu v podobě ukazatele  $R^2$ . Příkladem využitého regresního koeficientu pro GWR model je index stáří s nejvyššími hodnotami v jižní části Karlovarského kraje a na jižní Moravě. Naopak záporné hodnoty regresního koeficientu byly zjištěny na pomezí Jihomoravského kraje a kraje Vysočina.

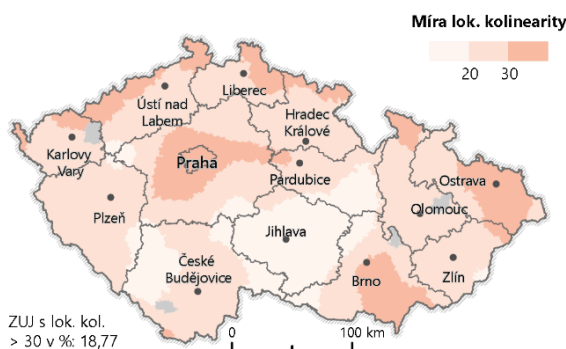
#### INDEX STÁŘÍ



#### PŘESNOST REGRESNÍHO MODELU



#### PROSTOROVÁ KOLINEARITA

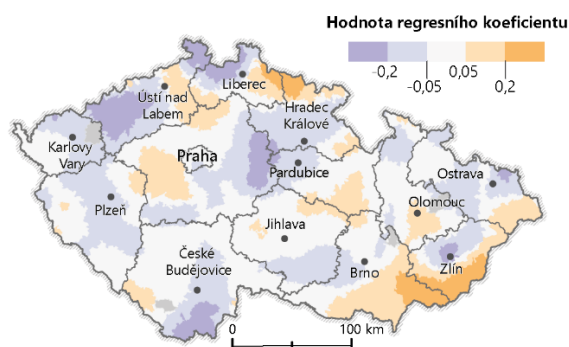


Obr. 36 Vybrané ukazatele GWR modelu pro hnutí ANO 2011

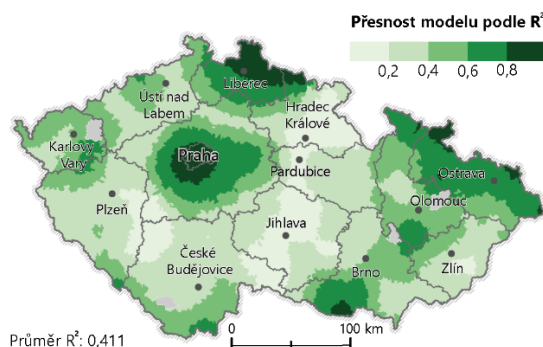
Z toho lze usoudit, že čím vyšší index stáří bude ve vymezené oblasti, tím nižší bude volební výsledek hnutí ANO 2011. Hodnotou ukazatele  $R^2$  je opět vykazována nejvyšší míra vysvětlení variability v okolí Prahy, nižší míra je naopak zaznamenána na západě Jihočeského kraje.

Třetí GWR model byl vytvořen za účelem vysvětlení výsledků koalice Pirátů a Starostů (Příloha 9). Výpočet byl proveden z pěti vstupních prediktorů. Zajímavostí u tohoto modelu je nejnižší naměřená míra lokální kolinearity ze všech vytvořených GWR modelů.

#### PODÍL PRIMÁRNÍHO SEKTORU



#### PŘESNOST REGRESNÍHO MODELU



#### PROSTOROVÁ KOLINEARITA

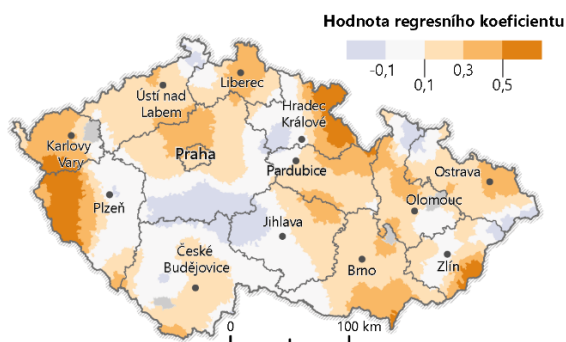


Obr. 37 Vybrané ukazatele GWR modelu pro koalici Pirátů a Starostů

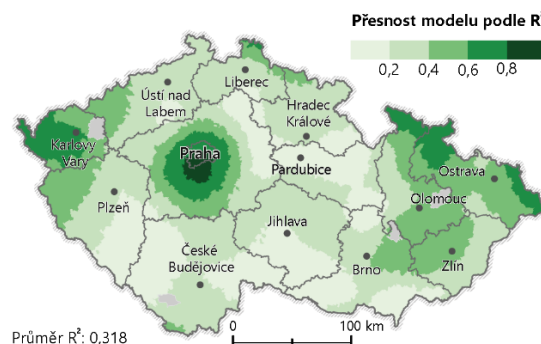
V GWR modelu koalice Pirátů a starostů je variabilita nejlépe vysvětlena kromě Prahy také na Liberecku a v příhraničních regionech Moravy a Slezska. Jedním z využitých prediktorů byl podíl osob zaměstnaných v primárním sektoru. V hodnotách regresního koeficientu převažují záporné hodnoty pro území ČR. Vyšší kladné hodnoty jsou viditelné pouze na jihovýchodě Moravy a na pomezí Libereckého a Královéhradeckého kraje. Podle vytvořeného modelu by vyšší hodnoty primárního sektoru v těchto oblastech byly následkem vyššího volebního zisku koalice Pirátů a Starostů. Nejvyšší míra vysvětlení variability je opět zaznamenána v Praze, ale v porovnání s ostatními modely také v okolí Liberce.

Poslední GWR model byl vypočítán z pěti vstupních prediktorů k vysvětlení volebního výsledku hnutí SPD jako závislé proměnné (Příloha 10). Ačkoliv je procentuální míra lokální kolinearity druhá nejnižší, z hlediska vysvětlení míry variability ukazatelem  $R^2$  lze vytvořený model označit za nejméně přesný ze všech vytvořených výstupů GWR. Příkladem regresního koeficientu je v tomto případě míra nezaměstnanosti, jež nabývá kladných hodnot u jihozápadních hranic ČR nebo v Královéhradeckém kraji. Záporné hodnoty tohoto koeficientu jsou viditelné například ve Šluknovském výběžku nebo na pomezí Středočeského kraje, kraje Vysočiny a Jihočeského kraje. Z výsledků regresního koeficientu lze opět odvodit kladné a záporné změny závislé proměnné v případě změny hodnoty prediktoru.

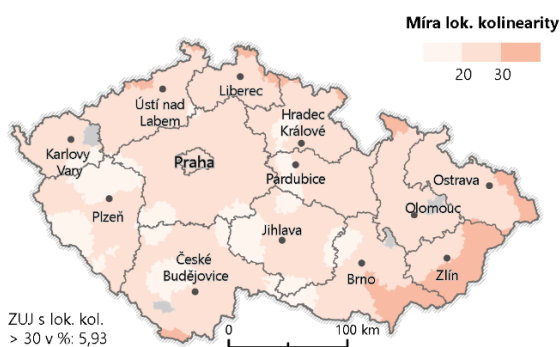
#### MÍRA NEZAMĚŠTANOSTI



#### PŘESNOST REGRESNÍHO MODELU



#### PROSTOROVÁ KOLINEARITA



Obr. 38 Vybrané ukazatele GWR modelu pro hnutí SPD

V případě nízké průměrné hodnotě  $R^2$  je viditelná vyšší přesnost modelu v okolí hlavního města Prahy nebo Jeseníků. V porovnání s ostatními GWR modely je ale vymezeno více oblastí s hodnotami nižšími než 0,2. Tento fakt může být způsoben zvolením méně vhodných prediktorů k vysvětlení volebního výsledku hnutí SPD. V rámci testovaných SED prediktorů ale nebyla nalezena kombinace prediktorů s vyšší mírou vysvětlení variability. Při analýze byla využita metoda GWR k vytvoření čtyř modelů, jimiž je vysvětlována míra variability závislé proměnné s využitím rozdílných prediktorů. Výsledky jednotlivých analýz jsou významným přínosem k objasnění vztahu volebního výsledku vybraných politických subjektů a použitých závislých proměnných v podobě prediktorů.

## 4.2.4 Predikce s využitím logistické regrese

V rámci čtvrté analýzy byla porovnána úspěšnost predikce s využitím klasických a prostorově vážených logistických modelů. Predikce na základě logistických modelů byla provedena pro oba kandidáty druhého kola prezidentských voleb 2023. Hlavní rozdíl logistického modelu oproti modelu lineární regrese spočívá ve využití sigmoidální funkce k modelování hodnot závislé proměnné binárního charakteru s využitím vstupních vysvětlujících proměnných (Kanade 2022a).

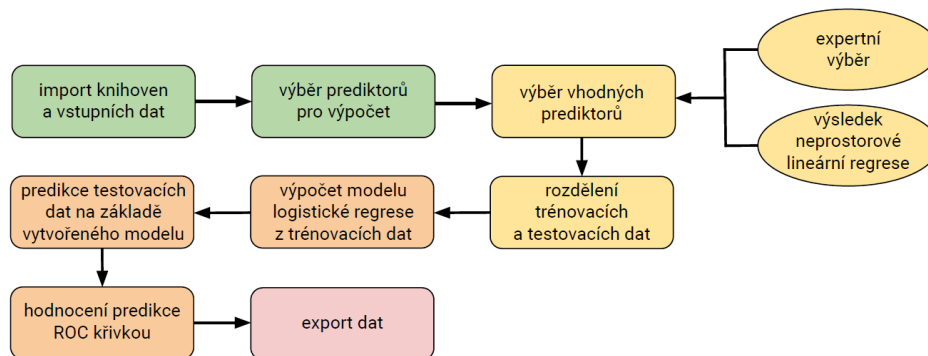
### Vytvoření logistických regresních modelů

Prvním krokem analýzy bylo vytvoření klasického neprostorového logistického modelu v jazyce R a prostorově váženého logistického modelu v ArcGIS Pro. Z důvodu porovnání úspěšnosti bylo využito stejných vstupních prediktorů pro oba druhy modelů. Jedinou nevyužitou metodou v kontextu hledání optimálních prediktorů byl aspekt prostorové významnosti. Důvodem byla absence tohoto vlivu u klasických regresních modelů. K vymezení optimálních vstupních prediktorů byly využity stejné metody jako při sestavení GWR modelů.

Tab. 7 Vstupní prediktory logistických modelů

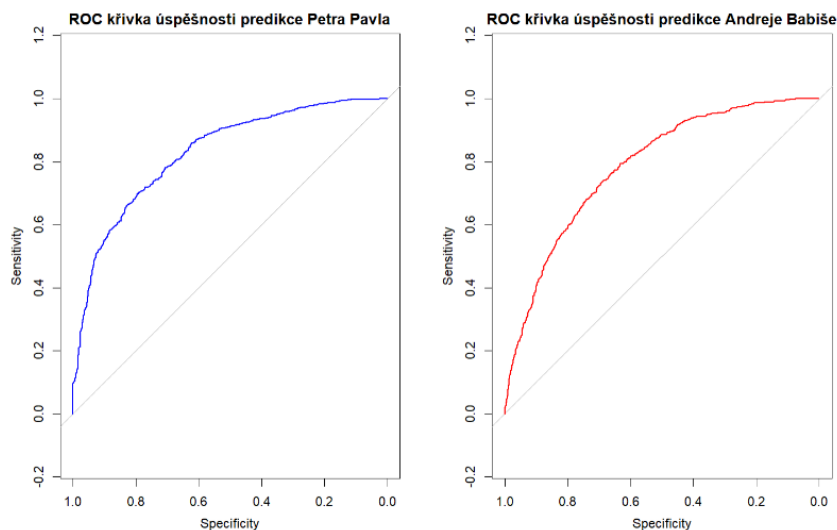
Vstupní atributy pro výpočet modelů logistické regrese	
Petr Pavel	Andrej Babiš
podíl zaměstnaných osob	index stáří
index kvality života	podíl nevěřících osob
podíl vysokoškolsky vzdělaných	podíl osob se středoškolským vzděláním bez maturity
podíl osob v exekuci	podíl osob s příspěvkem na bydlení
podíl osob s OB v jiném kraji než TB	podíl svobodných osob

Po výběru vstupních prediktorů byl napsán skript v jazyce R, jehož výstupem je predikce na základě klasického logistického modelu. Data byla nejprve rozdělena na trénovací data odpovídající jedné třetině záznamů a testovací data, jež byla tvořena zbylými dvěma třetinami záznamů. Tato technika je v rámci tvorby logistického regresního modelu aplikována, neboť se jedná o efektivní krok k vytvoření účinného modelu pro predikci. V opačném případě je velkým rizikem přeučení logistického modelu vedoucí až k nevyužitelnosti získaného modelu (Kanade 2022a). Dalším krokem bylo vytvoření klasického logistického modelu založeného na trénovacích datech. Následovala predikce hodnot na testovacích datech prostřednictvím vytvořeného logistického modelu. Posledním krokem před exportem predikovaných dat bylo hodnocení predikce modelu na základě ROC křivky.



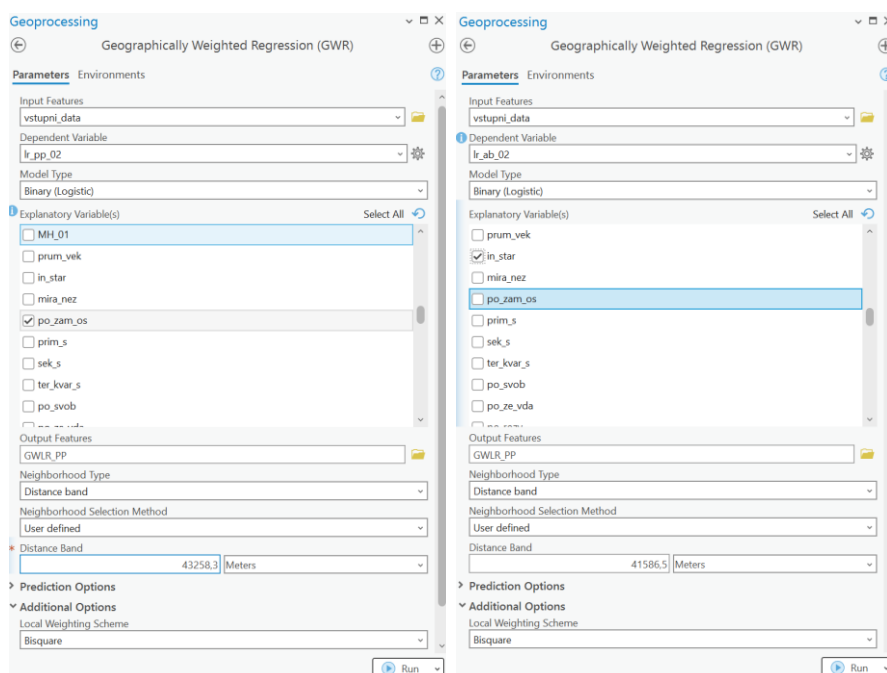
Obr. 39 Řešení pro predikci na základě logistického modelu

ROC křivka je technikou hodnocení úspěšnosti provedené predikce reprezentována grafem. V grafech je na svislé ose zobrazena senzitivita, což je míra schopnosti identifikace skutečně pozitivních případů. Naopak specificita je definována jako míra schopnosti identifikace skutečně negativních případů (Narkhede 2018). Hlavní částí grafu je tzv. AUC představující plochy pod červenou nebo modrou křivkou. Křivka AUC je hlavním indikátorem pro určení schopnosti správné klasifikace v rámci definovaných tříd. Z porovnání výsledků predikce pro Petra Pavla i Andreje Babiše byla zjištěna podobná úspěšnost u obou provedených predikcí.



Obr. 40 ROC křivka úspěšnosti predikce obou kandidátů na prezidenta

V druhém kroku tvorby byl vytvořen prostorově vážený logistický model nástrojem *Geographically Weighted Regression* v prostředí ArcGIS Pro. Toto prostředí bylo zvoleno z důvodu absence přístupu v balíku *GWmodel*. Jediným využitým nástrojem z tohoto balíku byl *bw.gwr* k určení optimální šířky vzdálenostního pásma prostorově váženého logistického modelu. K výpočtu prostorově váženého logistického regresního modelu byly využity stejné vstupní prediktory jako v případě klasických logistických regresních modelů.



Obr. 41 Vstupní nastavení výpočtu prostorově váženého logistického modelu v ArcGIS Pro

### Hodnocení úspěšnosti predikce logistických modelů

Po vytvoření klasického i prostorově váženého modelu bylo přistoupeno k hodnocení úspěšnosti predikce na základě druhu využitých logistických modelů. Za tímto účelem byly klasifikované hodnoty predikce zapsány do matice záměn, ve které jsou zaznamenány skutečně pozitivní a skutečně negativní případy vůči falešně pozitivním a negativním predikcím. Absolutní hodnoty bylo nutné přepočítat z důvodu porovnatelnosti na procentuální podíly z celku, neboť matice záměn klasického modelu je vytvářena pouze ze dvou třetin záznamů oproti datům z prostorového modelu.

MATICE ZÁMĚN Petr Pavel <i>klasický model</i>		Skutečné hodnoty		MATICE ZÁMĚN Andrej Babiš <i>klasický model</i>		Skutečné hodnoty	
		Pozitivní	Negativní			Pozitivní	Negativní
Predikované hodnoty	Pozitivní	2 471 58 %	696 16,3 %	Pozitivní	675 16 %	362 8,5 %	
	Negativní	326 7,8 %	766 17,9 %	Negativní	787 18,5 %	2 435 57 %	
MATICE ZÁMĚN Petr Pavel <i>prostorový model</i>		Skutečné hodnoty		MATICE ZÁMĚN Andrej Babiš <i>prostorový model</i>		Skutečné hodnoty	
		Pozitivní	Negativní			Pozitivní	Negativní
Predikované hodnoty	Pozitivní	3 816 59,7 %	729 11,4 %	Pozitivní	1 385 21,7 %	383 6 %	
	Negativní	393 6,2 %	1 450 22,7 %	Negativní	794 12,4 %	3 826 59,9 %	

Obr. 42 Matice záměn podle kandidátů a druhu logistických modelů

Následovalo zhotovení map, v nichž byl výsledek predikce zobrazen podle druhu logistických modelů. Za tímto účelem byly výsledky predikce spojeny a vytvořeny mapy podle klasických a prostorově vážených logistických modelů. Výsledná mapová kompozice je součástí diplomové práce (Příloha 11).

Z mapového výstupu bylo po další reklasifikaci hodnot zjištěno, že prostorově vážený logistický model byl z hlediska správnosti klasifikace vhodnější volbou než klasický logistický model. Úspěšnost byla hodnocena jako podíl úspěšné klasifikace v obou případech vůči ostatním případům, které nebyly úspěšné alespoň v jednom případě. Další skutečností je vysoká proporční podobnost jednotlivých hodnot u maticí záměn podle zvoleného kandidáta.

## 4.2.5 Shluková analýza

Pátá provedená analýza je věnována aplikaci metod a postupů k vymezení shluků z hlediska volebního chování. Prvním krokem bylo stanovení optimální metody a počtu výsledných shluků. Poté byla vybraná metoda aplikována a vymezené shluky byly následně charakterizovány volebními výsledky. Posledním krokem byl dodatečný popis shluků z hlediska SED ukazatelů.

### Výběr metody shlukování a stanovení počtu shluků

Nejprve bylo rozhodnuto o druhu využití metody shlukové analýzy. Po provedené rešerši bylo vybráno hierarchické shlukování. Hlavní výhodou hierarchického shlukování oproti nehierarchickému je využití konstruovaného dendrogramu k vymezení stanoveného počtu shluků (Šarmanová 2012). Hlavní přidanou hodnotou dendrogramu je sloučení či rozklad již existujících shluků k vytvoření shluků nových. Z tohoto důvodu je výsledek metody hierarchického shlukování snazší k interpretovat oproti nehierarchickému shlukování, kde jsou shluky odvozeny pouze rozkladem podmonožin s využitím vstupního algoritmu.

Po výběru druhu shlukování bylo přistoupeno k výběru samotné metody pomocí funkce *clValid* ze stejnojmenného balíku v jazyce R poskytujícího hodnocení metody z hlediska parametrů **Connectivity**, **Dunn** a **Silhouette**. Parametrem **Connectivity** je měřena míra spojení prvků zařazených do stejného shluku. **Dunn** parametrem je hodnoceno vzájemné oddělení jednotlivých shluků. Pomocí **Silhouette** je hodnocena kvalita provedeného shlukování porovnáním průměrných vzdáleností v rámci jednotlivých shluků (Brock a kol. 2021).

Následujícím krokem bylo vymezení počtu shluků. Za tímto účelem byla využita funkce *NbClust* k identifikaci optimálního počtu shluků podle zvolené metody. Hlavním výstupem funkce je graf, kterým jsou na svislé ose zobrazeny hodnoty **D-indexu**, kde dosaženým skóre je vyhodnocen počet shluků na vodorovné ose identifikací tzv. kolene grafu. **D-index** je v kontextu shlukové analýzy definována jako míra posouzení kvality počtu shluků zvolené metody a je založená na konceptu měření kompaktnosti shluků (Charrad a kol. 2013). Dále platí, že čím vyšší je hodnota **D-indexu**, tím slabší jsou vazby mezi prvky uvnitř výsledných shluků.

Hodnocení s využitím výše popsaných technik bylo provedeno u tří vybraných metod hierarchického shlukování. Před provedením samotného hodnocení byla zamítnuta metoda nejbližšího souseda, neboť při její aplikaci dochází k řetězení a výsledné nekompatnosti shluků. Další zavrženou metodou byla metoda centroidů, kterou jsou poskytovány nejlepší výsledky na datech s jasně oddělitelnými hranicemi (Šarmanová 2012).

První vybranou metodou byla shlukovací metoda McQuitty. Ta je v principu podobná metodě centroidů. Hlavním rozdílem je vyšší význam hranice mezi vytvořenými shluky, což je důsledkem lepší oddělitelnosti výsledných shluků v porovnání s metodou centroidů. Druhou testovanou metodou byla Wardova metoda s cílem minimalizace podobnosti shluků. Tento fakt je v kontextu analýzy volebních výsledků důležitý z hlediska vymezení unikátních regionů politických preferencí. Poslední testovanou metodou byla metoda průměrné vazby. Tato metoda je založena tvorbě shluků podle měření střední vzdálenosti mezi prvky v každém shluku. Vytvářené shluky jsou v rámci této metody v porovnání s ostatními metodami obvykle užší (Murphy 2021).

Z výsledků provedených analýz vyplývá, že nejlepších výsledků bylo dosaženo pomocí Wardovy metody shlukování. Hlavním rozdílem jsou vyšší vypočtené hodnoty parametru **Connectivity** vzhledem k metodě průměrné vazby nebo metodě McQuitty. Na základě tohoto zjištění lze konstatovat, že prvky shluků vymezených Wardovou metodou jsou propojeny silnějšími vazbami v porovnání s ostatními metodami. Ačkoliv byly

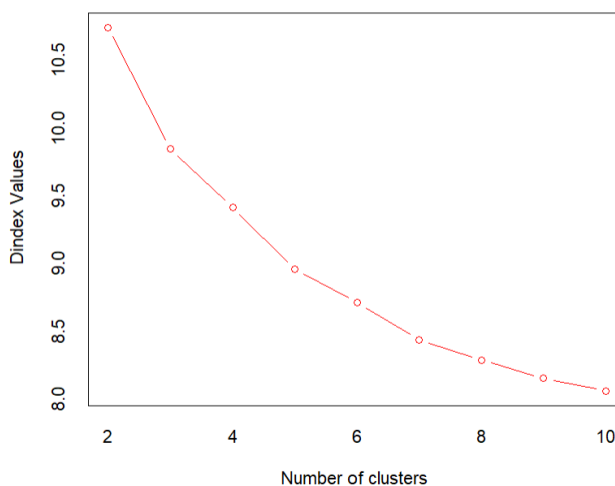


u parametrů **Silhouette** a **Dunn** zjištěny příznivější hodnoty metod McQuitty a průměrné vazby než Wardovy metody, výsledný rozdíl byl v porovnání s hodnotami parametru **Connectivity** zanedbatelný. Další zajímavostí je vyhodnocení dvou shluků jako optimálního vymezení u všech testovaných metod z hlediska vypočtených hodnot **Connectivity**, **Dunn** a **Silhouette**. Jedinou výjimkou byl výsledek **Dunn** parametru u Wardovy metody. Důsledkem tohoto zjištění je pravděpodobně volební preference určitých oblastí, např. Praha pro koalici SPOLU nebo Ústecký kraj pro ANO 2011. Na základě zjištěných poznatků byla pro provedení shlukové analýzy využita Wardova metoda.

Validation Measures:		2	3	4	5	6	7	8	9	10
hierarchical	Connectivity	8.2869	16.4377	41.9960	50.5250	50.5250	283.8083	344.4056	358.1282	386.6202
	Dunn	0.2258	0.1698	0.0762	0.0762	0.0762	0.0430	0.0446	0.0446	0.0458
	Silhouette	0.5415	0.5021	0.4427	0.4098	0.3882	0.2658	0.2422	0.2086	0.2015
Optimal Scores:										
	Score	Method	Clusters							
Connectivity	8.2869	hierarchical	2							
Dunn	0.2258	hierarchical	2							
Silhouette	0.5415	hierarchical	2							
<b>McQuittyho metoda</b>										
Validation Measures:		2	3	4	5	6	7	8	9	10
hierarchical	Connectivity	877.3266	1226.6278	1583.0865	1760.5976	2322.9163	2559.4433	2732.0083	2810.4992	2991.9198
	Dunn	0.0201	0.0202	0.0218	0.0218	0.0218	0.0218	0.0218	0.0218	0.0218
	Silhouette	0.2668	0.1908	0.1453	0.1393	0.0984	0.1058	0.0954	0.0854	0.0514
Optimal Scores:										
	Score	Method	Clusters							
Connectivity	877.3266	hierarchical	2							
Dunn	0.0218	hierarchical	4							
Silhouette	0.2668	hierarchical	2							
<b>Wardova metoda</b>										
Validation Measures:		2	3	4	5	6	7	8	9	10
hierarchical	Connectivity	8.2869	16.4377	41.9960	50.5250	50.5250	283.8083	344.4056	358.1282	386.6202
	Dunn	0.2258	0.1662	0.0762	0.0762	0.0762	0.0462	0.0446	0.0446	0.0446
	Silhouette	0.5415	0.5082	0.4427	0.4098	0.3882	0.2682	0.2422	0.2086	0.2049
Optimal Scores:										
	Score	Method	Clusters							
Connectivity	8.2869	hierarchical	2							
Dunn	0.2258	hierarchical	2							
Silhouette	0.5415	hierarchical	2							
<b>metoda průměrné vazby</b>										

Obr. 43 Hodnocení testovaných metod shlukové analýzy

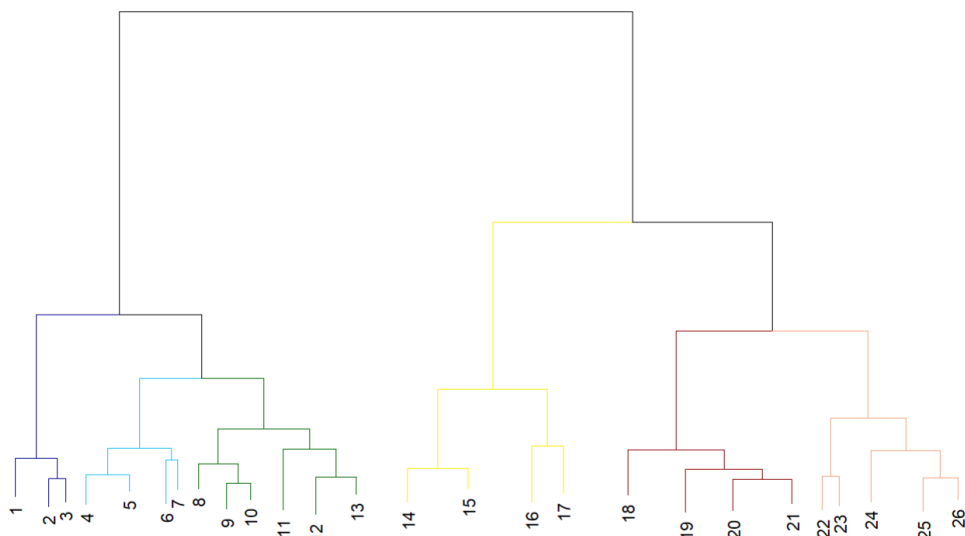
Určení optimálního počtu shluků bylo provedeno s využitím grafu počtu shluků, jenž byl vygenerován funkcí *NbClust* pro Wardovu metodu hierarchického shlukování. Jak již bylo dříve zmíněno, optimální počet shluků je obvykle identifikován jako tzv. koleno grafu. Po konzultaci s vedoucím práce bylo nakonec Wardovou metodou vymezeno šest výsledných shluků, ačkoliv bylo koleno grafu identifikováno na hodnotě sedm. Důvodem byla snažší budoucí interpretace šesti shluků z hlediska nižší variability prostorového rozmístění prvků.



Obr. 44 Míra kompaktnosti shluků Wardovy metody hierarchického shlukování

## Aplikace vybrané shlukovací metody

Po stanovení metody a počtu shluků bylo přistoupeno k aplikaci shlukovací metody na sedm politických subjektů s nejvyšším volebním výsledkem ve sněmovních volbách 2021. Nejprve byl vytvořen dendrogram ilustrující vývoj jednotlivých shluků a jejich vymezení. Jednotlivé větve dendrogramu jsou popsány čísly. Příkladem vazby mezi vymezenými shluky budiž azurový a zelený shluk, které by byly při určení nižšího počtu shluků spojeny v jeden celek. Z hlediska volebních výsledků je tento fakt způsoben vyšším procentuálním ziskem koalice Pirátů a Starostů v obou shlucích. Naopak je tomu u červeného a oranžového shluku, kde bylo nadprůměrných výsledků dosaženo hnutími ANO 2011 a SPD, které se nachází vůči současné vládě v opozici.



Obr. 45 Dendrogram vytvořených shluků

Následně bylo přistoupeno k charakterizaci vymezených shluků na základě volebních výsledků. Z důvodu vysokého nepoměru mezi subjekty s vyšším volebním ziskem oproti těm méně úspěšným bylo přistoupeno k charakterizaci shluků z hlediska rozdílu průměrných hodnot volebních výsledků daného shluku vůči celostátnímu průměru. Za tímto účelem byla sestavena tabulka rozdílu volebních výsledků vymezených shluků.

Tab. 8 Charakteristika vymezených shluků na základě volebních výsledků

	Rozdíly vůči průměrným hodnotám ČR u vybraných politických subjektů							
	SPOLU [%]	ANO 2011 [%]	Piráti a Starostové [%]	SPD [%]	Přísaha [%]	ČSSD [%]	KSČM [%]	
podpora vládní koalice	12,23	-8,16		1,62	-2,85	-0,34	-0,49	-1,38
podpora koalice Pirátů a starostů	-0,23	-4,29		6,39	-1,42	-0,16	-0,07	-0,36
podpora koalice Pirátů a Starostů a hnutí Přísaha	-6,06	-2,28		9,39	0,69	3,24	0,51	1,09
volební zisk odpovídající průměrným hodnotám ČR	2,25	0,74		-1,71	-0,73	-0,07	-0,16	-0,40
podpora vládní opozice	-11,89	14,23		-4,61	2,77	-0,48	-0,27	0,66
podpora hnutí SPD	-4,92	-1,35		-2,48	6,21	-0,05	0,89	0,94

První shluk byl na základě sestavené tabulky vyhodnocen jako shluk s nadprůměrnou podporou nynější vládní koalice SPOLU a Pirátů a Starostů. U koalice SPOLU byl zaznamenán nárůst volebního zisku o více než 12 %. Naopak je viditelný pokles volebního zisku u hnutí ANO 2011 o 8,16 % a SPD o 2,85 %. Typickými oblastmi tohoto shluku jsou Praha nebo okolí Brna.

Druhý shluk byl charakterizován nadprůměrným ziskem koalice Pirátů a Starostů o 6,39 %, podle kterého byl shluk pojmenován. Žádným jiným subjektem nebylo dosahováno v rámci shluku kladných hodnot a nejvyšší pokles byl opět zaznamenán u hnutí ANO 2011. Z prvků zařazených do tohoto shluku je tvořen téměř souvislý pás vedoucí od Prahy až po Liberecko.

U třetího vymezeného shluku byl zjištěn nejvyšší procentuální nárůst volebního zisku koalice Pirátů a Starostů. Vzhledem k celkovému volebnímu výsledku je tento shluk dále charakteristický vyšším volebním ziskem hnutí Přísaha. Kromě dvou zmíněných subjektů byl také identifikován drobný nárůst volebního výsledku u stran ČSSD a KSČM. Oblasti zařazené do třetího shluku lze vidět např. na Karlovarsku a v početné míře také na Ústecku.

Volební výsledky čtvrtého vymezeného shluku nejvíce odpovídaly průměru ČR, kdy u žádného z vybraných subjektů nebyl zaznamenán signifikantní pokles či nárůst volebního zisku. Zároveň se jedná o jediný shluk, kde nejúspěšnější politické subjekty, tedy koalice SPOLU a hnutí ANO 2011, zaznamenaly pozitivní nárůst volebního výsledku. Oblastí s vysokým výskytem prvků čtvrtého shluku je českomoravské pomezí v Pardubickém kraji a kraji Vysočina.

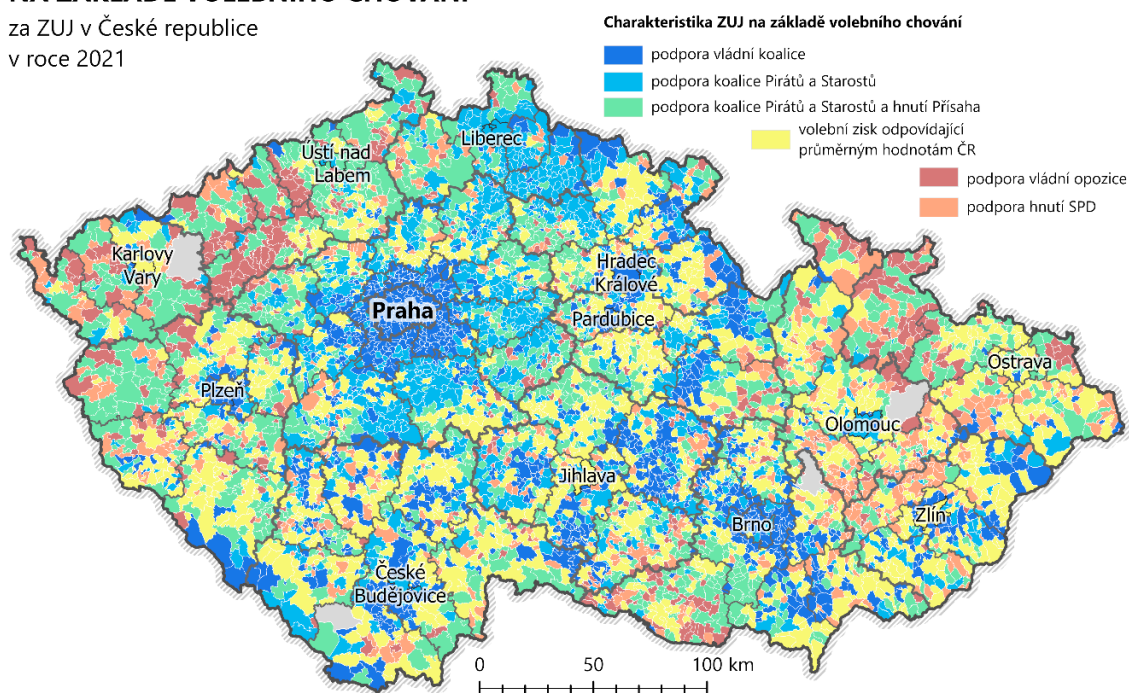
Pátý vymezený shluk byl na základě zjištěných rozdílů pojmenován jako shluk podpory vládní opozice. V tomto shluku byl viditelný nárůst volebního zisku pro hnutí ANO 2011 a SPD. Naopak tomu bylo u stran nynější vládní koalice s nejvyšší ztrátou volebního zisku v porovnání s ostatními shluky. Ve shluku jsou zahrnuty oblasti s nejvyšším procentuálním ziskem ANO 2011 identifikované při exploratorní analýze.

Pro šestý vymezený shluk byla charakteristická nadprůměrná podpora hnutí SPD. Dalším zjištěním byl nejvyšší procentuální nárůst u volebního výsledku ČSSD a částečně také u KSČM. Shluk je také typický poklesem volebních výsledků koalic SPOLU i Pirátů a Starostů. Prvky, zařazené do tohoto shluku, byly identifikovány například na pomezí Zlínského a Olomouckého kraje.

Z výsledku shlukové analýzy byla sestavena mapa, jejíž součástí byl i popis vymezených shluků z hlediska rozdílů volebních výsledků politických subjektů. Po provedení doplňkové charakteristiky shluků na základě SED ukazatelů byla do mapy přidána i tabulka s vybranými ukazateli, např. indexem stáří nebo indexem sociálního vyloučení. Mapa je součástí diplomové práce (Příloha 12).

## SHLUKY VYMEZENÉ NA ZÁKLADĚ VOLEBNÍHO CHOVÁNÍ

za ZUJ v České republice  
v roce 2021



Obr. 46 Mapa vymezených shluků volebního chování Wardovou metodou

### **Charakteristika shluků z hlediska SED ukazatelů**

Po vytvoření mapy výstupu shlukové analýzy byly vymezené shluky dále charakterizovány formou SED ukazatelů. V kontextu analýzy volebních výsledků je tato další charakteristika žádoucí, neboť je podkladem vysvětlení pro volebního chování.

Použité SED ukazatele jsou v tabulce rozděleny do tříd z hlediska tematického obsahu. K vzájemnému porovnání byly opět vypočteny rozdíly mezi průměrnými hodnotami ukazatelů vymezených shluků a celé ČR. Následně byly stanoveny intervaly k rozdělení hodnot využitých ukazatelů na podprůměrné, střední a nadprůměrné. Podprůměrné hodnoty byly definovány jako záporná změna alespoň o 15 % z rozsahu hodnot daného SED ukazatele vůči nulovému průměru. Rozsah hodnot byl určen hodnotami minima a maxima, kterých daný SED ukazatel nabýval. Naopak tomu bylo u nadprůměrných hodnot, kde musela být hodnota změněna alespoň o 15 % rozsahu hodnot v kladném směru. Zbývající hodnoty byly zahrnuty do kategorie středních hodnot. Definované intervaly byly barevně odlišeny. S pomocí výsledné tabulky byly sestaveny závěrečné interpretace shluků z hlediska SED charakteristik.

Shluk s převážnou podporou vládní koalice je nejprůzračnější z hlediska SED podmínek. Charakterizován je např. nadprůměrnou kvalitou života nebo vyšším procentuálním zastoupením vysokoškolsky vzdělaných osob. Dále se jedná především o městský region, neboť rozdíl hodnoty indexu vymezení města a venkova nabývá nejvyšších kladných hodnot. Z hlediska bydliště je zde nejvyšší podíl osob s trvalým bydlištěm v jiném kraji ČR nebo v zahraničí. Typickou obcí z hlediska SED ukazatelů je Holubov nacházející se v Jihočeském kraji.

Shluk nadprůměrné podpory koalice Pirátů a Starostů je charakteristický poměrně vysokým podílem žen se třemi dětmi a nejvyšším zastoupením osob pracujících v sekundárním sektoru. Nejtypičtější obcí tohoto shluku je Malešov ve Středočeském kraji. Charakteristickým rysem třetího shluku byl vyšší podíl ovdovělých osob a obecně starší obyvatelstvo v porovnání s průměrnými hodnotami. Z hlediska průměrných hodnot shluku je pro tento shluk charakteristická obec Dolní Město v kraji Vysočina.

Ve čtvrtém shluku, jehož volební výsledky byly nejbliže celostátnímu průměru, byla zjištěna nadprůměrná hodnota pouze u podílu osob s obvyklým bydlištěm v obci trvalého bydliště. Typickou obcí tohoto shluku je Luže v Pardubickém kraji.

Pátý vymezený shluk byl vyhodnocen z hlediska SED podmínek jako nejméně příznivý. Charakterizován je např. vysokou hodnotou indexu sociálního vyloučení nebo nejnižší hodnotou indexu kvality života. Charakteristickou obcí podle SED ukazatelů jsou Mikulovice v Olomouckém kraji.

V posledním vymezeném shluku s nadprůměrnou podporou SPD byl jako nejtypičtější obec vyhodnocen Velký Třebešov nacházející se v Královéhradeckém kraji. Typickou charakteristikou shluku je nadprůměrný podíl osob s maturitním vzděláním a vysoký podíl osob nehlásících se k žádné náboženské víře.

Tab. 9 Charakteristika vymezených shluků s využitím SED ukazatelů

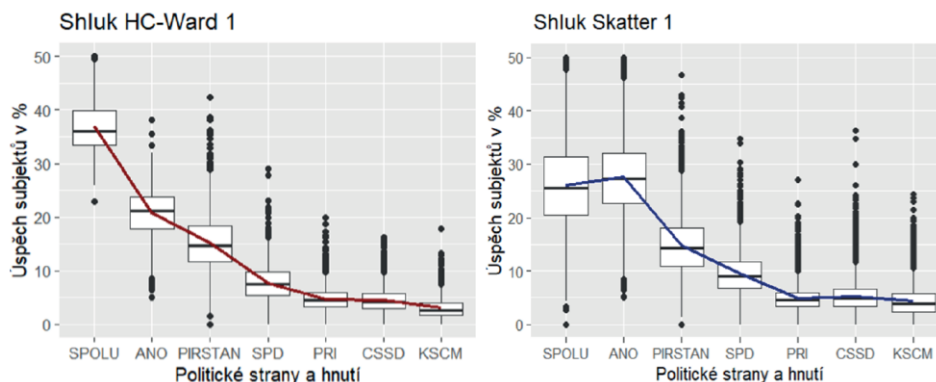
		Shluky vytvořené pomocí Wardovy metody hierarchického shlukování							
		podpora vládní koalice	podpora koalice Pirátů a starostů	podpora koalice Pirátů a Starostů a hnutí Přísaha	volební zisk odpovídající průměrným hodnotám ČR	podpora vládní opozice	podpora hnutí SPD	absolutní hodnota, odpovídající míře 15 % změny vzhledem k nulovému průměru	
Rozdíly mezi průměrnými hodnotami ČR a průměrnými hodnotami v jednotlivých shlucích	rodinný život	podíl svobodných [%]	-0,07	1,00	0,00	-0,69	1,25	-0,06	0,46
		podíl ženatých a vdaných [%]	1,90	-1,44	-0,98	0,83	-2,78	0,06	1,20
		podíl rozvedených [%]	-0,95	0,36	0,38	-0,29	1,05	0,28	0,50
		podíl ovdovělých [%]	-0,88	0,07	0,59	0,17	0,49	-0,29	0,37
		podíl žen bez dětí [%]	0,33	0,05	-0,18	-0,01	-0,11	-0,04	0,11
		podíl žen s 1 dítětem [%]	1,49	-1,13	-0,52	-0,34	-1,19	0,85	0,83
		podíl žen se 2 dětmi [%]	-0,19	-1,74	-0,80	1,24	-2,73	1,78	1,27
		podíl žen se 3 dětmi [%]	-1,64	2,82	1,49	-0,90	4,03	-2,59	2,02
	obvyklé a trvalé bydliště	podíl osob s OB stejným jako TB [%]	0,08	-0,21	0,29	2,03	-2,03	-2,00	1,00
		podíl osob s OB v okrese TB [%]	-2,89	1,94	0,24	0,93	0,99	-0,21	1,08
		podíl osob s OB v kraji TB [%]	-0,03	-0,43	0,13	0,07	0,18	-0,13	0,15
		podíl osob s OB v jiném kraji než TB [%]	2,76	-1,36	-0,85	-2,50	-0,12	2,32	1,49
		podíl osob s OB v zahraničí [%]	0,60	-0,18	-0,15	-0,40	0,37	0,14	0,27
	ekonomická aktivita a vzdělání	míra nez. [%]	-0,84	0,41	0,59	-0,37	1,70	-0,35	0,64
		primární s. [%]	-1,06	0,81	1,02	-0,96	3,18	-0,80	1,17
		sekundární s. [%]	-4,79	2,72	1,55	0,97	1,92	-0,83	1,92
		tercierní a kvartérní s. [%]	5,86	-3,53	-2,57	-0,01	-5,10	1,63	2,80
		podíl základního vzdělání [%]	-2,54	1,38	1,69	-0,71	4,39	-1,26	1,79
		podíl středního vzdělání bez maturity [%]	-5,68	3,08	2,77	-0,23	5,57	-1,67	2,85
		střední vzdělání s maturitou [%]	2,02	-1,70	-1,48	0,66	-4,36	1,82	1,81
		vysokoškolské vzdělání [%]	6,21	-2,76	-2,98	0,29	-5,59	1,11	2,84
		podíl nevěřících [%]	-3,16	-1,22	1,27	-1,84	3,51	3,45	2,17
	náboženství	podíl věřících Římskokatolické církve [%]	3,78	-0,28	-1,45	1,69	-3,42	-2,81	2,01
		index stáří [-]	-2,09	0,01	1,46	0,03	2,57	-0,75	1,04
		index kvality života [-]	4,35	-1,56	-2,86	1,48	-7,14	1,16	2,78
		index soc. vyloučení [-]	-1,00	0,00	-1,00	0,00	2,00	1,00	1
		index města a venkova [-]	0,77	-0,49	0,01	0,36	-0,97	-0,41	0,45
		počet obyvatel [-]	820	-915	-23	389	-994	-278	513
		Typická ZUJ	Holubov	Malešov	Dolní Město	Luže	Mikulovice	Velký Třebešov	
	ID ZUJ	545490	534188	568601	571776	540382	574589		

← -15 % změna      0      +15 % změna →  
nižší hodnota      střední hodnota      vyšší hodnota

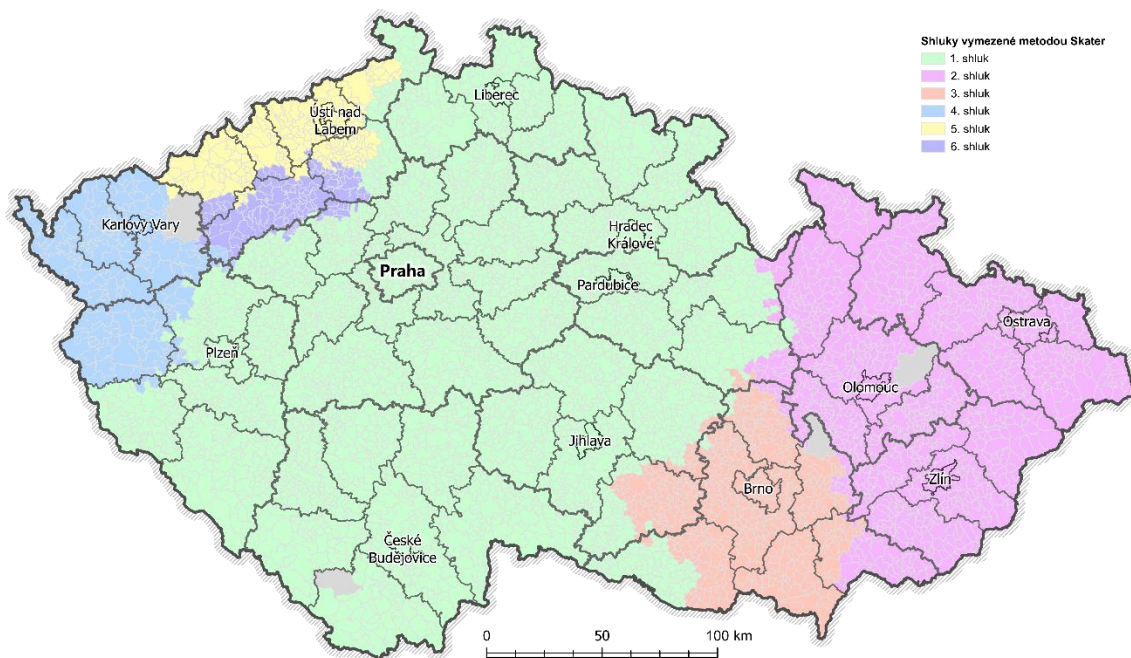
### Prostorově vážená shlukovací metoda

Za účelem porovnání výsledků klasického a prostorově váženého shlukování byla otestována také funkce *Skatter* z balíku *rGeoda*. Metodou *Skater* je v rámci vytváření shluků kombinována atributová podobnost s prostorovou podobností a výsledkem metody jsou prostorově kompaktnější shluky v porovnání s klasickými metodami (Anselin 2020b). V rámci provedené analýzy bylo vytvořeno šest výsledných shluků stejným způsobem, jako u klasického shlukování Wardovou metodou. Z hlediska prostorové kompaktnosti bylo dosaženo částečně příznivého výsledku, neboť shluky byl prostor rozdělen na homogenní části bez vnořených prvků jiného shluku. Nepříznivým zjištěním z prostorového hlediska bylo nerovnoměrné rozdělení prvků do jednotlivých shluků, protože jedním shlukem bylo pokryto téměř celé území Čech, zatímco v jiném shluku nebyl zahrnut ani počet prvků, odpovídající průměrné rozloze kraje ČR. Hlavní negativní výsledek

byl zaznamenán v atributovém hledisku metody z důvodu ztráty jasných rozdílů mezi nejméně úspěšnými subjekty, tedy koalici SPOLU a hnutím ANO 2011. Příkladem tohoto jevu je sloučení Prahy a okolí typické volbou koalice SPOLU s prvky zasahující na území sousedních krajů. Z tohoto důvodu nebyla metoda v rámci analýzy volebních výsledků dále využita.



Obr. 47 Porovnání hodnot shluku zahrnující Prahu podle využití metody



Obr. 48 Prostorové rozložení shluků vymezených metodou Skatter

## 5 VÝSLEDKY

Výsledkem diplomové práce je sestavená datová sada volebních výsledků obohacená o SED ukazatele a pět provedených analýz za účelem prostorového hodnocení dat. Tématem prostorových analýz byly volby do poslanecké sněmovny 2021 a prezidentské volby 2023. Hlavním výstupem analýz jsou vytvořené mapové výstupy obsahující kromě samotných map také nadstavbové kompoziční prvky, jako grafy, tabulky apod.

### 5.1 Sestavení datové sady

Dílčím cílem zpracované diplomové práce bylo sestavení datové sady z volně dostupných dat volebních výsledků (podkapitola 4.1). Hlavním datovým zdrojem volebních výsledků byl datový portál ČSÚ, ze kterého byla stažena data volebních výsledků a dalších ukazatelů, např. počet odevzdaných obálek nebo počet voličů v seznamu. Po stažení dat s prostorovou podrobností volebních okrsků bylo přistoupeno k agregaci volebních okrsků na ZUJ a následný přepočítání hodnot sloupců na procentuální podíly z důvodu stejné prostorové podrobnosti dat volebních výsledků a SED ukazatelů (podkapitola 4.1.1)

Datová sada byla následně obohacena o vybrané SED ukazatele (podkapitola 4.1.2), kde hlavním datovým zdrojem byl geoportál SLBD 2021, jenž je spravován ČSÚ. Z geoportálu SLBD 2021 byla stažena data s tematikou vzdělanosti, rodinného stavu, ekonomické aktivity apod. Prostorová podrobnost získaných dat odpovídá ZUJ. Dalším datovým zdrojem byla data využitá ke zpracování indexu sociálního vyloučení prostřednictvím MPSV. Příkladem jsou získaná data podílu osob v exekucích, podílu osob s příspěvkem na bydlení a samotný index sociálního vyloučení. Nakonec byly SED ukazatele obohaceny o index kvality života a index vymezení města a venkova. Data sociálního vyloučení a obou zmíněných indexů byla disgregována z prostorové podrobnosti odpovídající obcím na ZUJ nahrazením jednotné hodnoty daného ukazatele u větších územních celků. S výjimkou indexu kvality života, jenž byl získán pro rok 2023, odpovídají všechna data SED ukazatelů hodnotám za rok 2021. Podoba sestavené datové sady je zobrazena prostřednictvím tabulky volebních výsledků a SED ukazatelů (Tab. 1).

Po sestavení datové sady byla provedena exploratorní analýza (podkapitola 4.1.3), jejímž cílem bylo získání základních poznatků o volebním chování v České republice. Příkladem úlohy v rámci exploratorní analýzy je test normality dat pomocí K-S testu. Výsledkem provedených testů bylo zjištění o rozdělení dat, které s výjimkou dvou SED ukazatelů neodpovídalo normálnímu rozdělení (Obr. 12). Další využitou metodou byla prostorová autokorelace s využitím Moranova I, kterou bylo potvrzena tendence ke shlukování u vybraných ukazatelů volebních výsledků (Tab. 2) a vznik lokálních odlehlostí v rámci shluků vysokých či nízkých odlehlostí hodnot (Obr. 15). Poznatky zjištěné prostřednictvím exploratorní analýzy ovlivnily výběr prostorových statistických metod ke zpracování samotných analýz.

### 5.2 Aplikace prostorových analýz

Po sestavení datové sady bylo přistoupeno k aplikaci prostorových analýz (podkapitola 4.2). Prostorové analýzy byly zaměřeny na hodnocení mnohorozměrné odlehlosti, vysvětlení volebních výsledků na základě sestavených regresních modelů nebo klasifikace území na základě shlukové analýzy.

#### **Mahalanobisova vzdálenost**

První provedená analýza byla zaměřena na mnohorozměrné hodnocení odlehlosti s využitím Mahalanobisovy vzdálenosti (podkapitola 4.2.1). Nejprve byla hodnocena

odlehlost z globálního hlediska, kdy byly odlehle hodnoty zjišťovány na základě všech záznamů vstupního datového souboru. V rámci řešení byla vypočtena robustní verze Mahalanobisovy vzdálenosti, která je méně citlivá k odlehlým hodnotám, s využitím vymezených atributů sněmovních a prezidentských voleb (Tab. 4). Z tohoto důvodu byly jako odlehle hodnoty klasifikovány pouze nejvíce odlehle hodnoty, z hlediska stanovené prahové hodnoty odpovídající 95 % kvantilu Chí kvadrát rozdělení. Součástí analýzy jsou rovněž příklady klasifikované odlehlosti podle druhu voleb (Obr. 18, 19 a 20). Z vytvořeného mapového výstupu (Příloha 3) byl zjištěn vyšší výskyt odlehlých ZUJ při krajských hranicích. Krajem s nejvíce klasifikovanými odlehlými hodnotami byla Vysočina, nejméně odlehlých hodnot pak bylo zjištěno v Moravskoslezském kraji.

V rámci analýzy byla odlehlost hodnocena také z lokálního pohledu. Při lokálním pohledu byl výpočet Mahalanobisovy vzdálenosti proveden pouze pro prvky definovaného okolí, nikoliv pro celý rozsah datového souboru. K výpočtu byla v tomto případě využita klasická verze Mahalanobisovy vzdálenosti z důvodu vyšší citlivosti u klasifikace odlehlých hodnot. Při návrhu skriptu k výpočtu lokální Mahalanobisovy vzdálenosti byly identifikovány problémy s definicí okolí (Obr. 21) a vznikem singulární kovarianční matice (Obr. 22). Vzniklé problémy byly vyřešeny využitím nižšího rozsahu atributů pro výpočet (Tab. 5) a definicí okolí s dostatečným počtem prvků stanovených provedeným testováním na 220 nejbližších sousedů. K porovnání lokální a globální odlehlosti byla v tomto případě doložena i klasická verze globální Mahalanobisovy vzdálenosti. Prahová hodnota byla i v tomto případě stanovena na 95 % kvantilu Chí kvadrát rozdělení. Výstupem lokální analýzy odlehlosti jsou mapové výstupy hodnotící sněmovní volby 2021 (Příloha 4) a prezidentské volby 2023 (Příloha 5). Za účelem porovnání globální i lokální odlehlosti byly uvedeny identifikované příklady ZUJ (Obr. 26, 27 a 28). V obou vytvořených výstupech byl zjištěn vyšší výskyt odlehlých hodnot při sněmovních volbách 2021, který byl způsoben zahrnutím více atributů pro výpočet. Dalším faktem byl nižší počet lokálních odlehlostí v porovnání s globálními, a to v obou mapových výstupech. Z tohoto zjištění byl vyvozen závěr, že v lokálním definovaném okolí byla nižší tendence k výskytu ZUJ s atypickými volebními výsledky vůči globálnímu pohledu zahrnujícímu do výpočtu všechny prvky datového souboru.

Detekce odlehlých hodnot prostřednictvím Mahalanobisovy vzdálenosti je velice účinnou metodou k odhalení odlehlostí. Hlavní výhodou je především možnost výpočtu pro definovaný rozsah atributů, čímž je eliminována potřeba určení odlehlých hodnot odděleně v každém atributu. Sestavenými řešeními prostřednictvím skriptů v jazyce R (Přílohy 15 a 16) je umožněno provést klasifikaci odlehlosti z globálního i lokálního pohledu.

### **Analýza využití voličských průkazů**

Cílem druhé provedené analýzy bylo zmapování využití voličských průkazů při prezidentských volbách 2023 (podkapitola 4.2.2), kde byl předpokládán odlišný podíl voličů do lokalit významných lyžařských středisek. Prvním krokem této analýzy byl neúspěšný pokus vyhledání datové sady s evidencí voličských průkazů za ZUJ nebo jiné územní celky. Spolu s konzultantem práce v oblasti politologie byl proto navrhnut alternativní postup pro zjištění míry využití voličských průkazů spočívající v porovnání rozdílu voličů v seznamu mezi prezidentskými volbami 2023 a volbami do zastupitelstev krajů 2020. Za tímto účelem byl vypočten index změny získaný podílem voličů v seznamu v roce 2023 vůči voličům v seznamu uvedených k roku 2020. Signifikantní míra změny voličů v seznamu, kterou bylo indikováno nadprůměrné využití voličských průkazů, byla stanovena pro hodnoty 1,25 a vyšší. Signifikantní míra změny s hodnotou 1,25 je interpretována jako nárůst voličů v seznamu o 25 % v roce 2023 oproti roku 2020.



Mapový výstup porovnávající voliče v seznamu (Příloha 6) byl doplněn o sloupcové kartodiagramy, jejichž součástí byly kromě voličů v seznamu také hodnoty počtu obyvatel ZUJ se signifikantní mírou změny vypočteného indexu. Výsledkem analýzy využití voličských průkazů bylo 29 ZUJ signifikantní změny indexu a počtem voličů v seznamu v roce 2023 významně vyšším v porovnání s ostatními sloupci vytvořených kartodiagramů. Jedinou výjimkou byla obec Strážný v Jihočeském kraji (Obr. 31), kde počtem voličů v seznamu v roce prezidentských voleb nebyl překonán počet obyvatel dané obce, přestože byl překonán počet voličů v seznamu pro rok 2020. Z tohoto důvodu lze navrhnout více alternativních hypotéz pro vysvětlení změny, např. významná změna věkové struktury obyvatelstva obce.

Ačkoliv nebylo v provedené analýze počítáno přímo s daty využitých voličských průkazů, z vypočteného indexu byl zjištěn významný nárůst voličů v seznamu u obcí v horských oblastech blízko středisek zimních sportů, čímž byla potvrzena hypotéza o přesunu voličů mimo jejich obvyklá bydliště. Z tohoto důvodu je využitá metoda dobrou alternativou vysvětlující využití voličských průkazů při absenci dat požadované atributové a prostorové podrobnosti.

### **Geograficky vážená regrese**

Další provedená analýza byla věnována vytvoření GWR modelů k vysvětlení volebních výsledků čtyř politických subjektů ve sněmovních volbách 2021 jako závislých proměnných s nejvyšším volebním ziskem pomocí SED ukazatelů v podobě jejich prediktorů (podkapitola 4.2.3).

K přesnějším výsledkům GWR modelů byl využit balík *GWmodel*, jehož metodika byla popsána Golliniovou a kol. (2015). V sestaveném skriptu výpočtu GWR modelů (Obr. 32) byly využity funkce výpočtu optimálního vzdálenostního pásma funkcí *bw.gwr*, funkce *gwr.model.selection* k seřazení prediktorů podle prostorové významnosti, funkce *gwr.basic* pro vytvoření samotného GWR modelu a funkce *gwr.collin.diagno* zjišťující lokální kolinearitu vytvořeného GWR modelu. K výběru vhodných prediktorů byly využity také poznatky o vzájemných vztazích z exploratorní analýzy a ukazatele klasického modelu lineární regrese (Obr. 33). Následným krokem bylo vytvoření GWR modelů využívajících prediktory závislých proměnných, které byly stanoveny výše popsanými metodami (Tab. 6). Výsledkem provedené analýzy jsou prostorové regresní modely ve formě mapových kompozic pro koalici SPOLU (Příloha 7), hnutí ANO 2011 (Příloha 8), koalici Pirátů a Starostů (Příloha 9) a hnutí SPD (Příloha 10). Příklady vztahu regresních koeficientů, úspěšnosti vytvořeného GWR modelu ve vysvětlení variability závislé proměnné a hodnoty lokální kolinearit jsou prezentovány formou dílčích výstupů mapových kompozic (Obr. 35, 36, 37 a 38). Vytvořenými GWR modely jsou nastíněny komplexní vztahy mezi volebními výsledky vybraných politických subjektů a SED ukazatelů, vysvětlující např. pozitivní vztah míry nezaměstnanosti a volebního výsledku hnutí ANO 2011 na západě Ústeckého kraje prostřednictvím regresního koeficientu.

Metoda geograficky vážené regrese je vhodnou metodou k vysvětlení chování závislé proměnné v podobě volebních výsledků politických subjektů. Díky provedeným optimalizacím výsledného modelu je ve vytvořených modelech minimalizován efekt prostorové kolinearit, čímž bylo zamezeno kolizi prediktorů ve vysvětlení chování volebních výsledků.

### **Predikce s využitím logistické regrese**

Ve čtvrté provedené analýze byla zkoumána úspěšnost predikce na základě klasického a prostorově váženého modelu logistické regrese v kontextu druhého kole prezidentských voleb 2023 (podkapitola 4.2.4). Aplikací poznatků z exploratorní analýzy a výsledků

klasické lineární regrese byly stanoveny vhodné prediktory k vytvoření logistického modelu (Tab. 7).

V prvním kroku analýzy byla vytvořena predikce na základě klasických modelů logistické regrese. Za tímto účelem byla nejprve data rozdělena na trénovací a testovací. V dalším kroku byly z trénovacích dat vypočteny modely logistické regrese, na jejichž základě byla poté provedena predikce na testovacích datech. K hodnocení poměru senzitivity a specificity byly vygenerovány ROC křivky provedené predikce (Obr. 40). Následovalo vytvoření modelů prostorově vážené logistické regrese v prostředí ArcGIS Pro (Obr. 41), kde vstupní šířka vzdálenostního pásma byla vypočtena funkcí *bw.gwr* podobně, jako u GWR modelů v předchozí analýze.

Nakonec byly výsledky spojeny podle druhu využitého logistického modelu a vytvořeny matice záměn (Obr. 42) hodnotící přesnost provedených predikcí. Zkoumáním matic záměn bylo zjištěno, že prostorově vážené logistické modely byly přesnější než predikce na základě klasických modelů logistické regrese. Tento fakt je způsoben přidáním prostorového kontextu, jenž není u predikce klasických logistických modelů zahrnut. Výstupem analýzy predikce s využitím logistické regrese jsou mapy klasifikovaných predikcí podle správnosti (Příloha 11).

Logistické regresní modely byly shledány jako účinné v kontextu predikce volebních výsledků druhého kola prezidentských voleb. Z hlediska úspěšnosti predikce je lepší volbou prostorový logistický model oproti klasickému logistickému modelu. Nevýhodou této metody je potřeba dat binárního charakteru závislé proměnné k výpočtu logistické regrese, čímž může být zkomplikováno vytvoření obdobného modelu např. pro sněmovní volby, kde je předpokládána účast více než dvou politických subjektů.

### **Shluková analýza**

V poslední zpracované analýze je zkoumáno využití shlukové analýzy k vymezení oblastí z hlediska volebních výsledků sedmi nejúspěšnějších politických subjektů sněmovních voleb 2021 (podkapitola 4.2.5). V prvním kroku shlukové analýzy byla jako nejvhodnější identifikována Wardova metoda hierarchického shlukování, kde byla dále hodnocena využitelnost McQuittyho metody a metody průměrné vazby (Obr. 43). K hodnocení vhodnosti byly využity balíky *cValid* pro výběr metody a balík *NbClust* k hodnocení optimálního počtu shluků, který byl nakonec stanoven na šest.

Následně byla provedena aplikace hierarchického shlukování na data volebních výsledků a byl vygenerován dendrogram (Obr. 45), kterým je popsán proces slučování a rozkladu stanovených shluků. Další rozeznatelnou informací je vzájemná vzdálenost shluků. Příkladem budiž vzájemná blízkost shluku s identifikovanou nadprůměrnou podporou Pirátů a Starostů azurové barvy a shluku nadprůměrné podpory Pirátů a Starostů spolu s hnutím Přísaha zelené barvy. K vymezení z hlediska volebního chování byla vytvořena tabulka průměrných rozdílů volebních výsledků stanovených shluků vůči celostátnímu průměru (Tab. 8). S pomocí tabulky průměrných rozdílů byly stanovené shluky pojmenovány na základě nadprůměrných volebních zisků politických subjektů (např. shluk podpory vládní koalice). Za účelem dalšího popisu vymezených shluků bylo přistoupeno k charakteristice pomocí SED ukazatelů. Výstupem této dílčí analýzy je tabulka se třemi stanovenými intervaly vypočtených rozdílů SED ukazatelů jednotlivých shluků a celostátního průměru (Tab. 9).

Z výsledků shlukové analýzy byly identifikovány typické ZUJ, podporující vládní koalici či opozici, jimiž je možné doložit tvrzení z prvotního pohledu na data volebních výsledků u exploratorní analýzy (podkapitola 4.1.3). Z hlediska SED ukazatelů byl dále vyhodnocen shluk nadprůměrné podpory vládní koalice jako nejprůběžnější, a naopak shluk podporující

vládní opozici jako nejméně příznivý. Ostatní charakteristiky výsledných shluků jsou uvedeny v kapitolách Aplikace vybrané shlukovací metody a Charakteristika shluků z hlediska SED ukazatelů. Nakonec byl vytvořen mapový výstup obsahující mapu vymezených shluků a další nadstavbové kompoziční prvky (Příloha 12).

Z hlediska vymezení charakteristických shluků volebního chování byla použita Wardova metoda shlukové analýzy správnou volbou. Důkazem tohoto tvrzení je tabulka popisující rozdíly volebních výsledků vymezených shluků. Následnou charakteristikou s pomocí SED ukazatelů byly získány další informace o životních podmínkách.

V průběhu shlukové analýzy byla vyzkoušena také prostorově vážená shlukovací metoda *Skater* z balíku *rGeoda*. Hlavním účelem metody *Skater* bylo vytvoření prostorově kompaktnějších shluků oproti klasické Wardově metodě hierarchického shlukování. Ačkoliv byly shluky vytvořené touto metodou prostorově kompaktní, z hlediska hodnot volebních výsledků byly ztraceny klíčové charakteristiky popisující úspěšnost koalice SPOLU v okolí Prahy (Obr. 48), kvůli zahrnutí tohoto regionu do shluku, jímž byly pokryty také regiony s vyššími volebními výsledky ostatních politických subjektů. Z tohoto důvodu nebyla metoda *Skater* v rámci shlukové analýzy využita.

## 6 DISKUSE

V průběhu zpracování diplomové práce byla odhalena řada problémů a skutečností, jež bylo potřeba za účelem zdárného dokončení práce vyřešit. V této kapitole je obsažen výčet problémů i uskutečněná či navržená řešení.

Výskyt prvních problémů byl zaznamenán již při sestavení datové sady u dat volebních výsledků sněmovních voleb 2021. U dat sněmovních voleb 2021 bylo přítomno více záznamů pro jednu obec z důvodu evidence preferenčních hlasů kandidátů daného politického subjektu, čímž bylo zkomplikováno další zpracování dat. Ačkoliv byl tento problém zdárně vyřešen, podle autora práce by mohla být od ČSÚ poskytnuta také méně komplexní verze datového souboru s evidencí součtu hlasů jednotlivých politických subjektů a s jedním jedinečným záznamem pro daný volební okrsek.

Dalším vyřešeným problémem byla nejednotnost vstupních dat v kontextu prostorové podrobnosti volebních výsledků a SED ukazatelů. Prvotním záměrem práce bylo zpracování prostorových analýz za volební okrsky, neboť při zadání diplomové práce byla obdržena historická data volebních výsledků a SED ukazatelů pro rok 2011 s touto prostorovou podrobností. Z důvodu chybějící evidence dat SED ukazatelů za volební okrsky bylo přistoupeno k agregaci vstupních dat volebních výsledků do podoby ZUJ. Na popsání téma byla vedena diskuse také s konzultantem diplomové práce, kterým bylo popsáno získání datové sady volebních výsledků a SED ukazatelů pro rok 2011, jež byla vytvořena formou speciální zakázky ČSÚ. Obdobným způsobem by musela být získána také aktuální datová sada, což bylo z časových a finančních důvodů zamítnuto.

V rámci mnohorozměrného hodnocení odlehlosti Mahalanobisovou vzdáleností byly vzniklé problémy i řešení nastíněny již při samotném zpracování. U lokální Mahalanobisovy vzdálenosti se jednalo o okolí s nízkým počtem vstupních prvků a vznik singulární kovarianční matice, což bylo vyřešeno vymezením okolí s 220 nejbližšími sousedy a výběrem pouze části sloupců vstupujících do výpočtu. Z důvodů odlišných vstupních parametrů a použití klasické verze Mahalanobisovy vzdálenosti bylo znemožněno porovnání výsledků klasifikovaných lokálních s první zhotovenou mapou globálních odlehlostí, kde byla odlehlost určena na základě výpočtu robustní verze Mahalanobisovy vzdálenosti. Za účelem porovnání globálních a lokálních odlehlostí byla vypočtena klasická verze globální Mahalanobisovy vzdálenosti pro celý datový soubor se sloupci využitými při výpočtu lokální Mahalanobisovy vzdálenosti.

Ačkoliv bylo navrženým řešením dosaženo požadovaných výsledků, z pohledu autora diplomové práce je možné jej vylepšit odstraněním dílčích nedostatků. Prvním nedostatkem byl omezený výběr atributů k výpočtu, jež byl podmíněn zamezením vzniku singulární kovarianční matice znemožňující další výpočet. Dalším nedostatkem byl poměrně vysoký počet zahrnutých prvků definovaného okolí k výpočtu lokální Mahalanobisovy vzdálenosti. Tyto nedostatky byly způsobeny prostorovou variabilitou volebních výsledků, ačkoliv byl v určitých částech území ČR výpočet umožněn i s nižším počtem prvků okolí nebo více vstupními atributy. Z hlediska globálního porovnání lokálních odlehlostí by ale nebylo možné aplikovat postup s různými vstupními parametry. Odlišným případem by byl výzkum odlehlosti na základě různých nastavení vymezeného okolí či počtu atributů pro jednu nebo více konkrétních ZUJ. Výsledkem by pak mohla být zjištění závislosti klasifikace lokální odlehlosti dané ZUJ na vymezeném okolí ohraničeného intervaly v závislosti na počtu nejbližších sousedů. Dalším obohacením navrženého řešení by mohlo být přidání prostorových vah k prvkům vymezeného okolí, na jejichž základě by byl přisouzen vyšší význam bližším prvkům vůči prvkům vzdálenějším.

Ačkoliv je lokální podobě Mahalanobisovy vzdálenosti věnováno několik odborných publikací, příkladem budiž článek Filzmosera a kol. (2014), výsledná klasifikace odlehlosti je prováděna na základě přepočtených indexů odlehlosti, které jsou obvykle provedeny mimo geografický prostor v kartézském souřadnicovém systému. Z tohoto důvodu se autor práce domnívá, že navržené řešení výpočtu lokální Mahalanobisovy vzdálenosti může být přínosem k určování odlehlosti i v dalších prostorových analýzách.

Hlavním problémem analýzy využití voličských průkazů byla absence datové sady s touto tematikou. Za účelem alespoň částečného vysvětlení jejich využití byl proto s konzultantem práce v politologické oblasti navržen alternativní postup. Navržený postup byl složen z výpočtu indexu změny voličů v seznamu pro uvedené roky, následovaném doplněním sloupcových kartodiagramů o počty obyvatel v obcích se zjištěnou pozitivní mírou signifikantní změny pro rok 2023. Přestože podle výsledků navrženého postupu bylo identifikováno 29 ZUJ s vysoce pravděpodobným využitím voličských průkazů, stále jím nemůže být nahrazena práce s přímou evidencí využití voličských průkazů. Hlavním problémem je odhalení dalších ZUJ, kde byly voličské průkazy využity, ale z důvodu již zmíněné chybějící evidence je není možné navrženým postupem identifikovat. Chybějící evidence voličských průkazů za administrativní jednotky je také pravděpodobně příčinou, proč nebyla v rámci rešerše nalezena publikace zabývající se prostorovou či jinou kvantifikací voličských průkazů.

U analýzy popisu volebních výsledků s využitím GWR modelů bylo nejvíce času věnováno optimalizaci výpočtu modelu prostřednictvím volby vstupních prediktorů. GWR modely byly vytvořené podle metodiky popsané Golliniovou a kol. (2015), jejíž úspěšné otestování je podle autora práce hlavní přidanou hodnotou provedené analýzy. Využitím popsané metodiky bylo umožněno sestavení přesnějších GWR modelů prostřednictvím funkcí k výpočtu optimální šířky vzdálenostního pásma nebo seřazení prediktorů podle prostorové významnosti. Výsledkem využitých funkcí bylo dosaženo optimalizovaných GWR modelů z hlediska vysvětlení míry variability závislé proměnné spolu se zjištěnou mírou lokální kolinearity. Přestože je v prostředí ArcGIS Pro umožněn výpočet geograficky vážené regrese, zmíněné optimalizační nástroje zatím nebyly implementovány, což může vést kupříkladu k nižší míře vysvětlené variability vytvořeným GWR modelem.

V politologických analýzách byly GWR modely využity např. Lyskem a kol. (2020). Ani v této studii ale není uvedena informace o využití *GWmodel* balíku, jímž by mohly být vytvořené GWR modely optimalizovány. Důsledkem optimálnějších GWR modelů by mohla pak být rozdílná interpretace u využitých regresních koeficientů v rámci provedené studie.

V předposlední provedené analýze, zabývající se úspěšností predikce na základě klasického a prostorově váženého logistického modelu, je k diskusi zejména využitelnost tohoto zjištění v kontextu budoucího předpovídání volebních výsledků. Politické preference mohou být vinou vnějších či vnitřních jevů změněny, čímž by zcela jistě vznikly odlišnosti mezi predikovanými a skutečnými hodnotami. Zajímavé výsledky by mohly být předloženy predikcemi za více sněmovních či prezidentských voleb v průběhu let a následné hodnocení úspěšnosti predikcí v kontextu jednotlivých územních jednotek.

Poslední provedenou analýzou byly nejprve vymezeny shluky, které byly následně popsány na základě volebního výsledku a dodatečně charakterizovány vybranými SED ukazateli. Za účelem výběru shlukovací metody bylo provedeno hodnocení vhodnosti využitých metod pomocí balíků *clValid* a *NbClust*. Z provedených testování byly zjištěny nejlepší výsledky hodnotících parametrů u Wardovy metody hierarchického shlukování, jež byla následně k provedení analýzy využita. Z důvodu vyšší míry přehlednosti o možnostech rozkladu či spojení stávajících shluků byly upřednostněny hierarchické metody shlukování nad nehierarchickými. Z výsledků provedené shlukové analýzy byly vymezeny shluky

charakteristického volebního chování. Příkladem může být vymezený shluk vládní koalice či opozice. Shluky byly dále charakterizovány SED ukazateli, jimiž byly dokresleny další informace o míře vzájemných vztahů mezi volbou a SED podmínkami dané základní územní jednotky.

V porovnání se studií popsanou v článku Nurmiho a Meskanena (2018) bylo využito rozdílné metody k výpočtu shluků. Použitá metoda pro vymezení shluků v jihozápadním Finsku je založena na algoritmu Aleskerova a Nurmiho z roku 2008 využívající křivky relativních hodnot podpory čtyř politických stran. Wardova metoda hierarchického shlukování použitá ve shlukové analýze diplomové práce je naopak vypočtena ze sedmi relativních hodnot volebních výsledků politických subjektů pro území celého státu. Díky této skutečnosti je viditelný komplexní přehled vymezení shluků na celostátní úrovni. U výsledků shlukové analýzy byl dále diskutován s konzultantem práce čtvrtý vytvořený shluk, jehož volební výsledky byly nejbližší celostátnímu průměru. Podle konzultanta práce by tyto obce mohly být předmětem dalšího výzkumu s cílem porovnání výsledků ZUJ čtvrtého shluku s dalšími výsledky sněmovních voleb. Důvodem je změna voličských preferencí, pozorovaná v minulých volbách. Pokud by u vybraných ZUJ byla tato skutečnost prokázána, mohlo by se jednat o českou podobu tzv. swing states, které jsou významným faktorem pro výhru voleb v USA (Cohen 2023).

## 7 ZÁVĚR

Diplomová práce je zaměřena na hodnocení volebních výsledků z hlediska klasických prostorových statistických metod. Cílem práce bylo provedení analýz za účelem hodnocení odlehlosti, popisu volebního chování prostřednictvím regresních modelů a vymezení shluků podobných hodnot na sestavené datové sadě volebních výsledků a SED ukazatelů. Sestavená datová sada volebních výsledků a SED ukazatelů posloužila jako vstupní data pro analýzy sněmovních voleb 2021 a prezidentských voleb 2023. Datová sada je rozdělena na část volebních výsledků, v níž jsou obsaženy informace o výsledcích participujících politických subjektů, a dalších volebních ukazatelů, např. volební účasti nebo voličů v seznamu. Ve druhé části datové sady jsou uvedeny ukazatele popisující SED charakteristiky. Sestavená datová sada odpovídá po úpravách prostorové podrobnosti ZUJ. Prvotní poznatky k analýzám volebního chování byly získány v rámci exploratorní analýzy, jež byla provedena po sestavení datové sady.

V první provedené analýze byla hodnocena mnohorozměrná míra odlehlosti s využitím Mahalanobisovy vzdálenosti z globálního i lokálního pohledu u sněmovních voleb 2021 i prezidentských voleb 2023. Při globálním pohledu byla Mahalanobisova vzdálenost vypočtena pro všechny prvky datového souboru, zatímco u lokálního pohledu do výpočtu vstoupily pouze prvky okolí, vymezeného 220 nejbližšími sousedy. Odlehlost podle druhu voleb a metody výpočtu je demonstrována na příkladech s vysvětleným důvodem klasifikované odlehlosti. Při hodnocení bylo dále zjištěno, že kraj Vysočina obsahuje nejvíce ZUJ klasifikovaných jako odlehlé. Dalším faktem je vyšší výskyt odlehlých hodnot v příhraničních oblastech a oblastech vnitřních periferií na hranicích krajů. Metoda Mahalanobisovy vzdálenosti byla shledána přínosnou v kontextu detekce odlehlých hodnot především z důvodu možnosti využití více než jednoho atributu k provedení výpočtu.

V rámci druhé analýzy bylo zkoumáno využití voličských průkazů v kontextu prezidentských voleb 2023. Vzhledem k neexistující evidenci těchto dat byl společně s konzultantem diplomové práce navržen alternativní postup založený na indexu změny, vypočteného podílem voličů v seznamu roku 2023 a voličů v seznamu roku 2020.

Z výsledků provedené analýzy bylo u 29 ZUJ indikováno nadprůměrné využití voličských průkazů podložené pozitivní signifikantní změnou vypočteného indexu, v blízkosti významných lyžařských středisek. Ačkoliv nebyla v průběhu zpracování využita data voličských průkazů z důvodu jejich absence, aplikovaný postup představuje přijatelnou alternativu s ohledem na vstupní data, která byla pro diplomovou práci k dispozici.

Tématem třetí zpracované analýzy byl popis volebního chování čtyř nejúspěšnějších politických subjektů sněmovních voleb 2021 pomocí vybraných SED ukazatelů, které v této analýze sloužily jako prediktory, na vytvořených GWR modelech. Optimalizované GWR modely byly sestaveny s pomocí balíku funkcí *GWmodel*, jehož využití je věnována použitá metodika Golliniové a kol. (2015). Výsledkem analýzy jsou čtyři sestavené prostorově vážené regresní modely zvolených politických subjektů, jimiž byly koalice SPOLU, hnutí ANO 2011, koalice Pirátů a Starostů a hnutí SPD. Ve vytvořených regresních modelech jsou popsány vztahy mezi závislou proměnnou volebního výsledku daného subjektu a SED ukazatele, jako prediktoru prostřednictvím regresních koeficientů. Dále jsou v modelech zaznamenány informace o lokální kolinearitě a úspěšnosti modelu ve vysvětlení prostorové variability závislé proměnné. Nejvíce vstupních prediktorů bylo využito u modelu pro hnutí ANO 2011. GWR modelem s nejvyšší naměřenou lokální kolinearitou byl GWR model hnutí SPD. Nejvyšší přesnost vysvětlení prostorové variability

byla u všech modelů zaznamenána v okolí Prahy, a naopak nejméně byla lokální variabilita vysvětlena v Jihočeském kraji a na Vysočině.

Ačkoliv není metoda geograficky vážené regrese v kontextu hodnocení dat volebních výsledků novou metodou, použitou metodikou v rámci diplomové práce jsou popsány kroky k optimalizaci vytvořených modelů, což je hlavní přidanou hodnotou této práce v kontextu regresních analýz.

Čtvrtá zpracovaná analýza se zabývala porovnáním úspěšnosti predikcí v druhém kole prezidentských voleb 2023 založených na klasických a prostorově vážených modelech logistické regrese. Na základě vypočtených modelů byla provedena predikce, jejíž výsledky byly v dalším kroku hodnoceny podle druhu využitého modelu logistické regrese. Z hodnocení prostřednictvím matice záměn bylo zjištěno, že prostorově vážený logistický model byl v predikci hodnot úspěšnější než klasický logistický model. Tato skutečnost je způsobena zahrnutím prostorového kontextu do výpočtu modelu logistické regrese.

Logistická regrese byla shledána účinným nástrojem k predikci volebních výsledků pro druhé kolo prezidentských voleb. Největší komplikací v tomto případě je potřeba dat binárního charakteru k sestavení modelu, čímž může být zkomplikováno provedení obdobné predikce pro jiný druh voleb. Důvodem je především vyšší počet participujících politických subjektů.

Poslední analýza byla zpracována za účelem vymezení shluků z hlediska volebních výsledků sedmi nejúspěšnějších politických subjektů ve sněmovních volbách 2021, jež byly dále charakterizovány prostřednictvím vybraných SED ukazatelů. Nejprve byly na základě teoretických charakteristik vybrány tři shlukovací metody, jež byly následně otestovány pomocí funkcí balíků *clValid* obsahujícího parametry, podle kterých byly vybrané metody hodnoceny, a *NbCLust* k určení optimálního počtu shluků. Z testovaných metod byla vybrána Wardova metoda hierarchického shlukování, jejíž aplikací bylo vymezeno šest výsledných shluků. U vymezených shluků byly v dalším kroku vypočteny rozdíly průměrných hodnot volebních výsledků jednotlivých shluků vůči celostátnímu průměru a sestavena tabulka, podle které byly výsledné shluky pojmenovány. K dodatečné charakteristice shluků byla dále vytvořena tabulka vybraných SED ukazatelů, kde byly opět vypočítány rozdíly mezi průměrnými hodnotami shluků a celostátním průměrem. Výsledné rozdíly SED ukazatelů v dalším kroku rozděleny na podprůměrné, střední a nadprůměrné hodnoty.

Použitá Wardova metoda hierarchického shlukování byla v kontextu vymezení shluků volebních výsledků zhodnocena jako velmi účinná. Důvodem jsou především vypočítané rozdíly průměrných hodnot shluků vůči celostátnímu průměru, čímž byly odhaleny preference volebního chování uvnitř shluků. Využití SED ukazatelů k dodatečné charakteristice poskytlo další informace o životních podmínkách, jimiž jsou ovlivněny preference voličů daných shluků.

Výstupy jednotlivých analýz jsou zpracovány formou desíti mapových výstupů, ve kterých jsou kromě samotných map uvedeny i další klíčové informace v podobě grafů, tabulek apod. Výsledky diplomové práce mohou posloužit nejen politologům nebo geografům, ale i politickým stranám a dalším subjektům, které mají zájem o výzkum volebních výsledků v prostorovém kontextu.



## POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

ANSELIN, Luc, 2020a. Local Spatial Autocorrelation. GeoDa [online] [vid. 2024-04-13]. Dostupné z: [https://geodacenter.github.io/workbook/6a\\_local\\_auto/lab6a.html](https://geodacenter.github.io/workbook/6a_local_auto/lab6a.html)

ANSELIN, Luc, 2020b. Spatial Clustering. GeoDa [online] [vid. 2024-04-24]. Dostupné z: [https://geodacenter.github.io/workbook/9c\\_spatial3/lab9c.html#skater-worked-example](https://geodacenter.github.io/workbook/9c_spatial3/lab9c.html#skater-worked-example)

AGGARWAL, Charu C., 2016. Outlier Analysis. 2. B.m.: Springer. ISBN 978-3-319-47577-6.

BAHOUNKOVÁ, Petra, 2013. VOLBY 2013: ČSSD vyhrála, důvodů k radosti ale moc nemá. ČT24 [online] [vid. 2024-04-29]. Dostupné z: <https://ct24.ceskatelevize.cz/clanek/domaci/volby-2013-cssd-vyhrala-duvodu-k-radosti-ale-moc-nema-317729>

BARTELS, Larry M., 2000. Partisanship and Voting Behavior, 1952-1996. American Journal of Political Science [online] [vid. 2024-04-21]. 35–50. Dostupné z: doi: 10.2307/2669291

BERNARD, Josef, Tomáš KOSTELECKÝ, 2014. Prostorový Kontext Volebního Chování – Jak Působí Lokální a Regionální Prostředí Na Rozhodování Voličů. Sociologický Časopis / Czech Sociological Review [online] [vid. 2024-04-21]. 50, 3–28. Dostupné z: doi: 10.13060/00380288.2014.50.1.30.

BIBEN, Martin, 2023. V „moravském Vatikánu“ dominoval Fischer, teď jsou místní nakloněni Pavlovi. Věřicím jsem něco naznačil, říká farář. Hospodářské noviny [online] [vid. 2024-04-21]. Dostupné z: <https://archiv.hn.cz/c1-67162280-v-bdquo-moravskem-vatikanu-ldquo-dominoval-fischer-ted-jsou-dolni-bojanovice-naklonene-pavlovi>

BLAŽEK, Tomáš, 2011. Rozdělování obcí na volební obvody: Gerrymandering v Česku? In: Cofola 2011 [online] [vid. 2024-04-21]. Brno: Tribun EU. ISBN 978-80-210-5582-7. Dostupné z: [https://www.law.muni.cz/sborniky/cofola2011/files/participace/Blazek\\_Tomas\\_6233.pdf](https://www.law.muni.cz/sborniky/cofola2011/files/participace/Blazek_Tomas_6233.pdf)

BOBBIT, Zach, 2021. What is Moran's I? Statology [online] [vid. 2024-04-14]. Dostupné z: <https://www.statology.org/morans-i/>

BROCK, Guy, Vasyl PIHUR, Susmita DATTA, Somnath DATTA, 2021. clValid, an R package for cluster validation [online] [vid. 2024-04-21]. 2021. B.m.: Department of Bioinformatics and Biostatistics, University of Louisville. Dostupné z: <https://cran.r-project.org/web/packages/clValid/vignettes/clValid.pdf>

CANSIZ, Sergen, 2021. Multivariate Outlier Detection in Python. Towards Data Science [online] [vid. 2024-04-14]. Dostupné z: <https://towardsdatascience.com/multivariate-outlier-detection-in-python-e946cfc843b3>

CHARRAD, Malika, Nadia GHAZZALI, Véronique BOITEAU, Azam NIKNAFS 2013. A examination of indices for determining the number of clusters : NbClust Package [online] [vid. 2024-04-21]. Dostupné z: [https://www.researchgate.net/publication/257138057\\_An\\_examination\\_of\\_indices\\_for\\_determining\\_the\\_number\\_of\\_clusters\\_NbClust\\_Package/citation/download](https://www.researchgate.net/publication/257138057_An_examination_of_indices_for_determining_the_number_of_clusters_NbClust_Package/citation/download)

CHARVÁT, Jakub, 2014. První přímá volba prezidenta ČR v roce 2013. Praha: Metropolitan University Prague Press. ISBN 978-80-86855-96-7.

COHEN, Jared, 2023. The rise of geopolitical swing states. Goldman Sachs [online] [vid. 2024-04-26]. Dostupné z: <https://www.goldmansachs.com/intelligence/pages/the-rise-of-geopolitical-swing-states.html>

COLE, Shawn, Andrew HEALY, Eric WERKER, 2012. Do voters demand responsive governments? Evidence from Indian disaster relief. Journal of Development Economics [online] [vid. 2024-04-21]. 97(2), 167–181. Dostupné z: doi: 10.1016/j.jdeveco.2011.05.005

ČSÚ, 2015. Metodika – volby [online] [vid. 2024-04-16]. Dostupné z: <https://www.czso.cz/csu/xs/metodika-volby>

ČSÚ, 2021. Sčítání lidu, domů a bytů 2021. SLBD 2021 [online] [vid. 2024-04-21]. Dostupné z: <https://geodata.statistika.cz/as/atlas/?xmax=1970309.4507756562&ymin=6304404.940638514&wkid=102100&ds=14916219&xmin=1461789.189000171&ymin=6304404.940638514&wkid=102100&ds=1>

ČSÚ, 2023. První kolo volby prezidenta provázela rekordní účast voličů. Katalog produktů [online] [vid. 2024-04-21]. Dostupné z: <https://www.czso.cz/csu/czso/prvni-kolo-volby-prezidenta-provazela-rekordni-ucast-volicu>

DOLEJŠÍ, Milan, 2018. Je sečteno. Zeman vyhrál, dostal 51 procent hlasů. ČT24 [online] [vid. 2024-04-21]. Dostupné z: <https://ct24.ceskatelevize.cz/tema/autori-webu-ct24-84343>

DOUBRAVOVÁ, Barbora, Anna DOHNALOVÁ, 2021. Desetina voličů váhá mezi koalicemi. „Spolu má hlasy jistější, Piráti musí zabojovat". Aktuálně.cz [online] [vid. 2024-05-06]. Dostupné z: <https://zpravy.aktualne.cz/domaci/volby/volby-nerozhodnuti-volici-spolu-pirati-stan/r~2dc34b86211511ec98380cc47ab5f122/>

DUIGNAN, Brian, 2024. gerrymandering [online] [vid. 2024-04-21]. B.m.: Britannica. Dostupné z: <https://www.britannica.com/topic/gerrymandering>

ESRI, 2024a. How Spatial Autocorrelation (Global Moran's I) works. Spatial Statistics [online] [vid. 2024-04-18]. Dostupné z: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-how-spatial-autocorrelation-moran-s-i-spatial-st.htm>

ESRI, 2024b. Incremental Spatial Autocorrelation. Spatial Statistics [online] [vid. 2024-04-18]. Dostupné z: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/incremental-spatial-autocorrelation.html>

EUROSTAT, 2021. Glossary:Herfindahl Hirschman Index (HHI) [online] [vid. 2024-04-21]. Dostupné z: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Herfindahl\\_Hirschman\\_Index\\_\(HHI\)&oldid=542901](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Herfindahl_Hirschman_Index_(HHI)&oldid=542901)

FILZMOSE, Peter, Anne RUIZ-GAZEN, Christine THOMAS-AGNAN 2014. Identification of local multivariate outliers. Statistical Papers [online] [vid. 2024-04-21]. 55(1). Dostupné z: [doi:10.1007/s00362-013-0524-z](https://doi.org/10.1007/s00362-013-0524-z)

FORCINA, Antonio, Giovanni M. MARCHETTI, 2011. The Brown and Payne Model of Voter Transition Revisited [online] [vid. 2024-04-21]. Dostupné z: [doi:10.1007/978-3-642-11363-5\\_54](https://doi.org/10.1007/978-3-642-11363-5_54)

GANTI, Akhilesh, 2024. Degrees of Freedom in Statistics Explained: Formula and Example. investopedia [online] [vid. 2024-04-21]. Dostupné z: <https://www.investopedia.com/terms/d/degrees-of-freedom.asp>

GOLLINI, Isabela, Binbin LU, Martin CHARLTON, Christopher BRUNSDON, 2015. GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models. Journal of Statistical Software [online] [vid. 2024-04-21]. 63(17), 1–50. Dostupné z: [doi:10.18637/jss.v063.i17](https://doi.org/10.18637/jss.v063.i17)

HÁJEK, Lukáš, Kamil ŠVEC, 2017. No Title. ČT24 [online] [vid. 2024-04-29]. Dostupné z: <https://ct24.ceskatelevize.cz/clanek/domaci/roztristena-snemovna-je-vyhodou-pro-anoroste-i-riziko-obstrukci-90279>

HIGDON, Dave, Jenise SWALL, John KERN, 1999. Non-Stationary Spatial Modeling. Bayesian statistics [online] [vid. 2024-04-21]. 6(1), 761–768. Dostupné z: [doi:10.48550/arXiv.2212.08043](https://doi.org/10.48550/arXiv.2212.08043)

INTERVIEW PLUS, 2019. Právce vs. levice? Ne, je to trojúhelník – liberalismus, konzervatismus a socialismus, vysvětluje politolog. Český rozhlas [online] [vid. 2024-05-06]. Dostupné z: <https://plus.rozhlas.cz/pravice-vs-levice-ne-je-trojuhelnik-liberalismus-konzervatismus-a-socialismus-8024463>

KANADE, Vijay, 2022a. What Is Logistic Regression? Equation, Assumptions, Types, and Best Practices. Spiceworks [online] [vid. 2024-04-13]. Dostupné z: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-logistic-regression/>

KANADE, Vijay, 2022b. What Is Spatial Analysis? Definition, Working, and Examples. Spiceworks [online] [vid. 2024-04-26]. Dostupné z: <https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-spatial-analysis/>

KOUBA, Karel, 2007. Prostorová analýza českého stranického systému. Institucionalizace a prostorové režimy. Czech Sociological Review [online] [vid. 2024-04-21]. 43(5), 1017–1038. Dostupné z: doi:10.13060/00380288.2007.43.5.08

KULACHAI, Waiphot, Unisa LERDTOMORNSAKUL, Patipol HOMYAMYEN, 2023. Factors Influencing Voting Decision: A Comprehensive Literature Review. Social Sciences [online] [vid. 2024-04-21]. 12, 469. Dostupné z: doi: 10.3390/socsci12090469

LI, Xu , Songren DENG , Lifang LI, Yunchuan JIANG, 2019. Outlier Detection Based on Robust Mahalanobis Distance and Its Application. Open Journal of Statistics [online] [vid. 2024-04-21]. 9(1), 15–26. Dostupné z: doi: 10.4236/ojs.2019.91002

LINKE, Andrew, John O' LOUGHLIN, 2016. Spatial Analysis in Political Geography. The Wiley Blackwell Companion to Political Geography [online] [vid. 2024-04-21]. 187–205. Dostupné z: doi:10.1002/9781118725771.ch15

LOPATECKI, Jason, 2023. Kolmogorov Smirnov Test for AI: When and Where To Use It [online] [vid. 2024-04-17]. Dostupné z: <https://arize.com/blog-course/kolmogorov-smirnov-test/>

LYSEK, Jakub, Jiří PÁNEK, Tomáš LEBEDA, 2020. Who are the voters and where are they? Using spatial statistics to analyse voting patterns in the parliamentary elections of the Czech Republic. Journal of Maps [online] [vid. 2024-04-21]. 17(1), 33–38. Dostupné z: doi: 10.1080/17445647.2020.1819901

LYSEK, Jakub, Karel MACKŮ, 2022. Continuity in Discontinuity? Spatial Patterns of Electoral Behaviour in the 2021 Czech Parliamentary Elections. Czech Journal of Political Science [online] [vid. 2024-04-21]. 3. Dostupné z: doi:10.5817/PC2022-3-260

MAISEL, L. Sandy, 2016. Who are Republicans? Who are Democrats? Who are the “others”? American Political Parties and Elections: A Very Short Introduction [online] [vid. 2024-04-21]. 2. vydání. Dostupné z: doi: 10.1093/actrade/9780190458164.003.0004

MÉSZÁROS, József, Norbert SOLYMOSI, Ferenc SPEISER, 2007. Spatial distribution of political parties in Hungary 1990–2006. Political Geography [online] [vid. 2024-04-21]. 26(7), 804–823. Dostupné z: doi: 10.1016/j.polgeo.2007.06.002

MINISTERSTVO VNITRA ČESKÉ REPUBLIKY, 2024. Volby v ČR [online] [vid. 2024-04-21]. Dostupné z: <https://www.mvcr.cz/volby/clanek/volby-v-cr.aspx>

MORAVEC, Zdeněk, 2018. Lineární regrese [online] [vid. 2024-04-21]. Dostupné z: <http://z-moravec.net/chemie/zaklady-chemie/linearni-regrese/>

MURPHY, Phil, 2021. Clustering Data in R [online] [vid. 2024-04-21]. Dostupné z: <https://rpubs.com/pjmurphy/599072>

NÁDVORNÍKOVÁ, Iveta, 2023. Chalupářka Nerudová. Když vyhraje volby, zahájí v Telecím hasičský ples. Svitavský Deník [online] [vid. 2024-04-21]. Dostupné z: [https://svitavsky.denik.cz/zpravy\\_region/chaluparka-nerudova-kdyz-vyhraje-volby-zahaji-v-telecim-hasiccky-ples-20230113.html](https://svitavsky.denik.cz/zpravy_region/chaluparka-nerudova-kdyz-vyhraje-volby-zahaji-v-telecim-hasiccky-ples-20230113.html)

NARKHEDE, Sarang, 2018. Understanding AUC - ROC Curve. Towards Data Science [online] [vid. 2024-04-23]. Dostupné z: <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>

NOVINKY.CZ, 2023. Volební účast: Data a historické srovnání pro všechny volby. *Novinky.cz* [online] [vid. 2024-05-04]. Dostupné z: <https://www.novinky.cz/tag/volebni-ucast-54700>

NURMI, Lasse Juhani, Tommi MESKANEN, 2007. Political areas of competition in Southwest Finland: A cluster analysis of parliamentary election data from three levels of aggregation. *Finnish Journal of Social Research* [online] [vid. 2024-04-21]. 11, 24–43. Dostupné z: [doi:10.51815/fjsr.110778](https://doi.org/10.51815/fjsr.110778)

PÁSZTO, Vít, Alžběta BRYCHTOVÁ, Pavel TUČEK, Lukáš MAREK, Jaroslav BURIAN, 2010. Using a fuzzy inference system to delimit rural and urban municipalities in the Czech republic in 2010. *Journal of Maps* [online] [vid. 2024-04-21]. 11(2), 231–239. Dostupné z: [doi: 10.1080/17445647.2014.944942](https://doi.org/10.1080/17445647.2014.944942)

PIRÁTI A STAROSTOVÉ, 2023. Co od nás můžete čekat? Tady je celý náš program. Piráti a Starostové [online] [vid. 2024-06-05]. Dostupné z: <https://www.piratiastarostove.cz/program/>

POLÁKOVÁ-UVÍROVÁ, Petra, 2024. Ministr Rakušan míří na Hanou se svou debatou bez cenzury. *Olomoucký deník* [online] [vid. 2024-04-21]. Dostupné z: [https://olomoucky.denik.cz/zpravy\\_region/do-prerova-miri-vit-rakusan-se-svou-debatou-bez-cenzury-20240306.html](https://olomoucky.denik.cz/zpravy_region/do-prerova-miri-vit-rakusan-se-svou-debatou-bez-cenzury-20240306.html)

PRABHAKARAN, Selva, 2024. Mahalanobis Distance – Understanding the math with examples (python) [online] [vid. 2024-04-21]. Dostupné z: <https://www.machinelearningplus.com/statistics/mahalanobis-distance/>

RAJABZADEH, Hossein, Mansoor Zolghadri JAHROMI, Mohammad Sadegh ZARE, Mostafa FAKHRAHMAD, 2018. Local Distance Metric Learning for the Nearest Neighbor Algorithm [online] [vid. 2024-04-23]. 2018. Shiraz: Department of Computer Science and Engineering, Shiraz University. Dostupné z: <https://arxiv.org/ftp/arxiv/papers/1803/1803.01562.pdf>

RAMACHANDRAN, Gowri, Dara GOLD, 2018. Using Outlier Analysis to Detect Partisan Gerrymanders: A Survey of Current Approaches and Future Directions. Election Law Journal: Rules, Politics, and Policy [online] [vid. 2024-04-21]. 17(4). Dostupné z: doi: 10.1089/elj.2018.0503

REY, Sergio J., 2001. Mathematical Models in Geography [online] [vid. 2024-04-21]. B.m.: San Diego State University, San Diego, California, USA. Dostupné z: doi: 10.1016/B0-08-043076-7/02516-X

RUSCH, Thomas, Ilro LEE, Kurt HORNIK, Wolfgang JANK, Achim ZEILEIS, 2013. Influencing elections with statistics: Targeting voters with logistic regression trees. The Annals of Applied Statistics [online] [vid. 2024-04-21]. 7(3), 1612–1639. Dostupné z: doi:10.1214/13-AOAS648

SCHOBBER, Patrick, Christa BOER, Lothar A. SCHWARTE, 2018. Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia [online]. 126(5), 1763–1768 [vid. 2024-04-17]. Dostupné z: doi:10.1213/ANE.0000000000002864

STEM, 2021. Stručná analýza letošního volebního chování [online] [vid. 2024-04-21]. Dostupné z: <https://www.stem.cz/strucna-analyza-letosniho-volebniho-chovani/>

STRMISKA, Zdeněk, 2017. Levice – pravice [online]. B.m.: Sociologický ústav AV ČR, v.v.i. Dostupné z: [https://encyklopedie.soc.cas.cz/w/Levice\\_-\\_pravice](https://encyklopedie.soc.cas.cz/w/Levice_-_pravice)

ŠARMANOVÁ, Jana, 2012. Metody analýzy dat. První vydání. Ostrava: VŠB – Technická univerzita Ostrava. ISBN 978-80-248-2565-6.

ŠIMÍČEK, Petr, 2014. Československo od revoluce k rozdělení státu. Od pádu komunismu k jednotné Evropě [online] [vid. 2024-04-28]. Dostupné z: <https://www.moderni-dejiny.cz/clanek/ceskoslovensko-od-revoluce-k-rozdeleni-statu/>

ŠIROKÝ, Jan, 2011. Tvoříme a publikujeme odborné texty. Praha: Computer Press. ISBN 978-80-251-3510-5.

ŠVEC, Kamil, Veronika MALÁ, 2018. Dohoda, která vymazala rozdíl mezi pravíci a levíci. Před 20 lety politici podepsali opoziční smlouvu. ČT24 [online] [vid. 2024-04-28]. Dostupné z: <https://ct24.ceskatelevize.cz/clanek/domaci/dohoda-ktera-vymazala-rozdil-mez-pravici-a-levici-pred-20-lety-politici-podepsali-opozicni-smlouvu-79008>

THE ECONOMIST, 2018. Imperial borders still shape politics in Poland and Romania [online] [vid. 2024-04-24]. Dostupné z: <https://www.economist.com/graphic-detail/2018/11/21/imperial-borders-still-shape-politics-in-poland-and-romania>

TODOROV, Valentin, 2024. Scalable Robust Estimators with High Breakdown Point - Robust Location and Scatter Estimation and Robust Multivariate Analysis with High Breakdown Point [online] [vid. 2024-04-21]. 2024. Dostupné z: <https://github.com/valentint/rrcov>

VLÁDA ČR, 2010. Přímá volba prezidenta [online] [vid. 2024-04-21]. Dostupné z: <https://vlada.gov.cz/cz/media-centrum/aktualne/prima-volba-prezidenta--78534>

VLÁDA ČR, 2024. Přehled vlád ČR [online] [vid. 2024-04-29]. Dostupné z: <https://vlada.gov.cz/cz/clenove-vlady/historie-minulych-vlad/prehled-vlad-cr/1993-2007-cr/>

WHEELER, David, Antonio PÁEZ, 2009. Geographically Weighted Regression [online]. B.m.: Elsevier Science [vid. 2024-04-14]. Dostupné z: doi: 10.1016/B978-008044910-4.00447-8

WIATR, Jerzy J., 1962. Economic and Social Factors of Electoral Behaviour. The Polish Sociological Bulletin [online] [vid. 2024-04-21]. 3, 65–75. Dostupné z: <https://www.jstor.org/stable/45275033>

WOLCHIK, Sharon, 1993. The Repluralization of Politics in Czechoslovakia. Communist and Post-Communist Studies [online] [vid. 2024-04-21]. 26(4), 412–431. Dostupné z: doi: 10.1016/0967-067X(93)90030-U

ZARYCKI, Tomasz, 2015. The electoral geography of Poland: between stable spatial structures and their changing interpretations. Erdkunde [online] [vid. 2024-04-21]. 69(2), 107–124. Dostupné z: doi: 10.3112/erdkunde.2015.02.02

# **PŘÍLOHY**



# SEZNAM PŘÍLOH

Kapitola se nečísluje. Uvést kompletní číslovaný seznam vázaných a volných příloh.

Příloha 1 (vázaná):

Korelační matice: Hodnoty Spearmanova korelačního koeficientu  
pro politické subjekty

Příloha 2 (vázaná)

Korelační matice: Hodnoty Spearmanova korelačního koeficientu  
pro SED ukazatele

Příloha 3 (vázaná a elektronická)

Mapa: Globální odlehle hodnoty v obou volbách  
za ZUJ v České republice v letech 2021 a 2023 (elektronická i volná)

Příloha 4 (vázaná a elektronická)

Mapa: Odlehle hodnoty sněmovních voleb  
za ZUJ v České republice v roce 2021 (elektronická i volná)

Příloha 5 (vázaná a elektronická)

Mapa: Odlehle hodnoty prezidentských voleb  
za ZUJ v České republice v roce 2023 (elektronická i volná)

Příloha 6 (vázaná a elektronická)

Mapa: Porovnání voličstva mezi volbami  
za ZUJ v České republice v letech 2020 a 2023 (elektronická i volná)

Příloha 7 (vázaná a elektronická)

Mapa: Prostorový regresní model koalice SPOLU  
v České republice v roce 2021 (elektronická i volná)

Příloha 8 (vázaná a elektronická)

Mapa: Prostorový regresní model hnutí ANO 2011  
v České republice v roce 2021 (elektronická i volná)

Příloha 9 (vázaná a elektronická)

Mapa: Prostorový regresní model koalice Pirátů a Starostů  
v České republice v roce 2021 (elektronická i volná)

Příloha 10 (vázaná a elektronická)

Mapa: Prostorový regresní model hnutí SPD  
v České republice v roce 2021 (elektronická i volná)

Příloha 11 (vázaná a elektronická)

Mapa: Porovnání predikce modelů logistické regrese  
za ZUJ v České republice v letech 2021 a 2023 (elektronická i volná)

Příloha 12 (vázaná a elektronická)

Mapa: Charakteristika a preference shluků  
za ZUJ v České republice v roce 2021 (elektronická i volná)

Příloha 13 (elektronická)

Skript: Skript\_01\_seskupeni\_pro\_snemovni\_volby

Příloha 14 (elektronická)

Skript: Skript\_02\_exploratorni\_analyza

Příloha 15 (elektronická)

Skript: Skript\_03\_globalni\_Mahalanobisova\_vzdalenost

Příloha 16 (elektronická)

Skript: Skript\_04\_lokalni\_Mahalanobisova\_vzdalenost

Příloha 17 (elektronická)

Skript: Skript\_05\_GWR\_modely

Příloha 18 (elektronická)

Skript: Skript\_06\_logisticka\_regrese

Příloha 19 (elektronická)

Skript: Skript\_07\_shlukova\_analyza

Příloha 20 (volná)

Poster

### **Popis struktury odevzdávaných digitálních dat**

Poster

Skripty

MapoveVystupy

Text\_Prace (text formát MS Word a PDF)

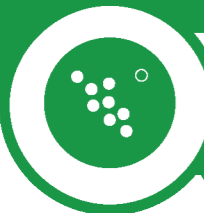
VstupniData

VystupniData

WEB







# GLOBALNÍ ODLEHLÉ HODNOTY V OBOU VOLBÁCH

za ZUJ v České republice v letech 2021 a 2023

Pro výpočet Mahalanobisovi vzdálenosti byla využita funkce **Mahalanobis** za použití **robustní kovarianční matice**.

Globální Mahalanobisova vzdálenost byla vypočtena z **6 388 prvků**.

Pro prezidentské volby bylo nastaveno **8** stupní volnosti a pro sněmovní volby bylo nastaveno **12** stupňů volnosti.

Pro prahovou hodnotu byl použit **95% kvantil Chí kvadrát**.

Atributy pro výpočet Mahalanobisovi vzdálenosti					
sněmovní volby	SPOLU	ANO	Piráti a Starostové	SPD	Přísaha
	ČSSD	Volný Blok	Zelení	KSČM	TSS
	ostatní politické strany, hnutí a koalice				
prezidentské volby	P. Pavel	A. Babiš	D. Nerudová	P. Fisher	
	J. Bašta	M. Hilšer	K. Diviš	T. Zima	
jiné	volební účast sněm. voleb		volební účast prez. voleb		

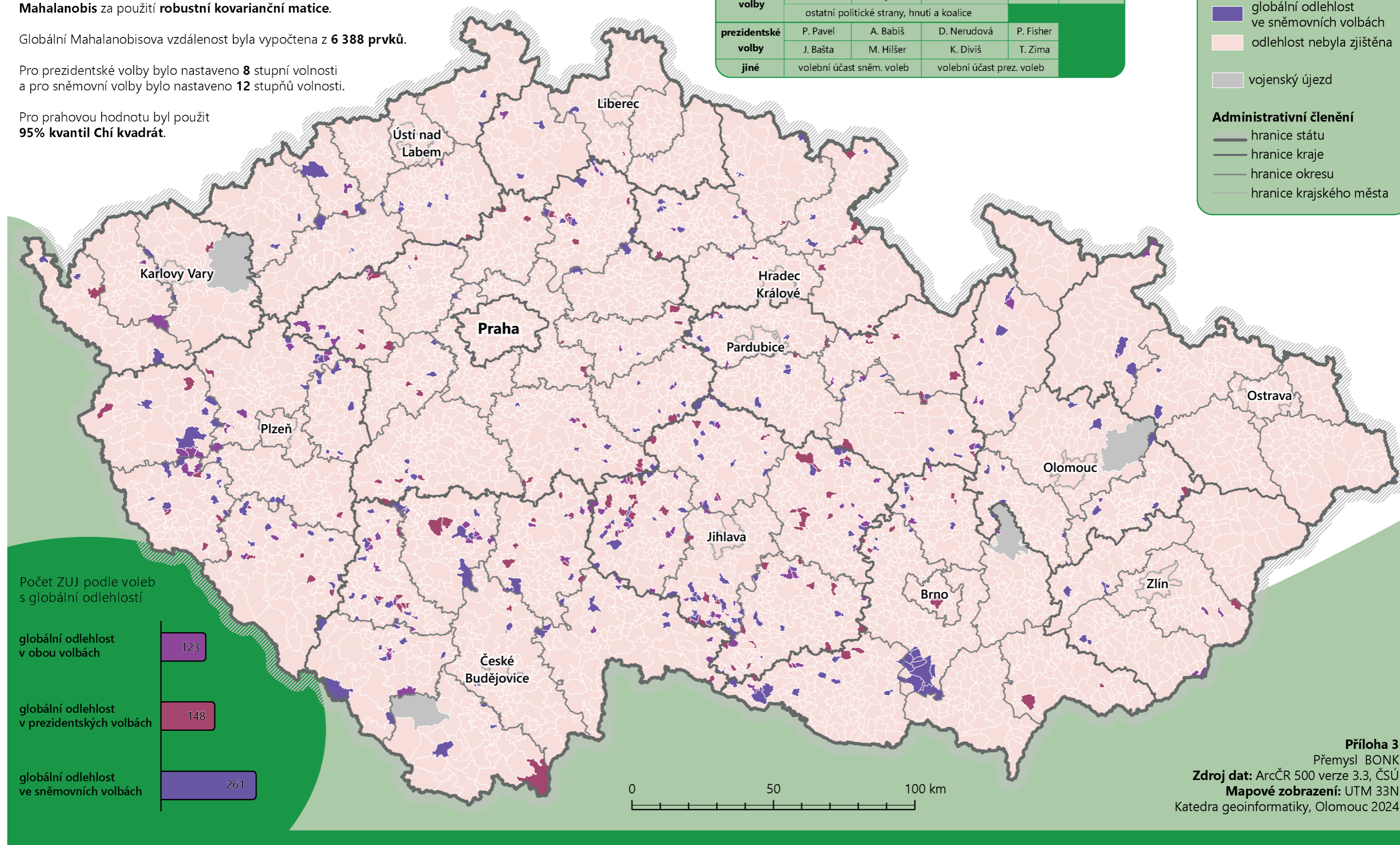
## Klasifikace odlehlosti pro ZUJ

- globální odlehlost v obou volbách
- globální odlehlost v prezidentských volbách
- globální odlehlost ve sněmovních volbách
- odlehlost nebyla zjištěna

vojenský újezd

## Administrativní členění

- hranice státu
- hranice kraje
- hranice okresu
- hranice krajského města



Počet ZUJ podle voleb s globální odlehlostí

globální odlehlost v obou volbách

123

globální odlehlost v prezidentských volbách

148

globální odlehlost ve sněmovních volbách

261

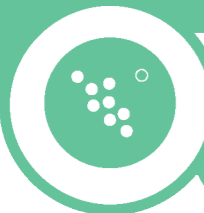
Příloha 3

Přemysl BONK

Zdroj dat: ArcČR 500 verze 3.3, ČSÚ

Mapové zobrazení: UTM 33N

Katedra geoinformatiky, Olomouc 2024



# ODLEHLÉ HODNOTY SNĚMOVNÍCH VOLEB

za ZUJ v České republice v roce 2021

Pro výpočet Mahalanobisovi vzdálenosti byla využita funkce **Moutlier**.

Globální Mahalanobisova vzdálenost byla vypočtena z **6 388 prvků**.

Lokální Mahalanobisova vzdálenost byla vypočtena z **220 prvků** zvolených metodou **nejbližšího souseda** každý prvek.

Pro prahovou hodnotu byl použit **95% kvantil Chí kvadrát**.

Atributy pro výpočet Mahalanobisovi vzdálenosti						
zisk > 5%	SPOLU	ANO	Piráti a Starostové	SPD		
zisk < 5%	Přísaha	ČSSD	KSČM	TSS	Volný Blok	Zelení
jiné	volební účast sněm. voleb					

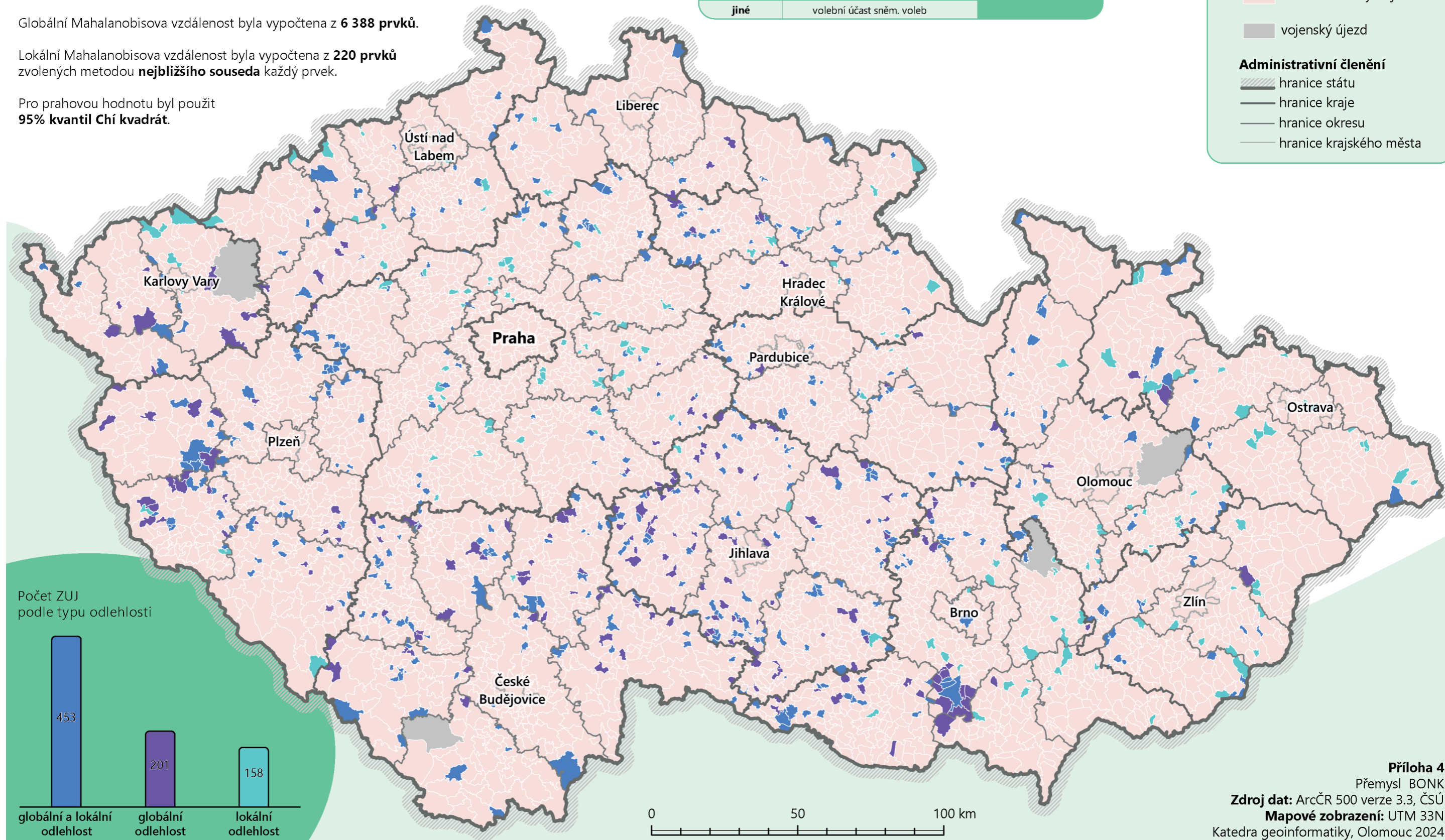
## Klasifikace odlehlosti pro ZUJ

- globální i lokální odlehlost
- globální odlehlost
- lokální odlehlost
- odlehlost nebyla zjištěna

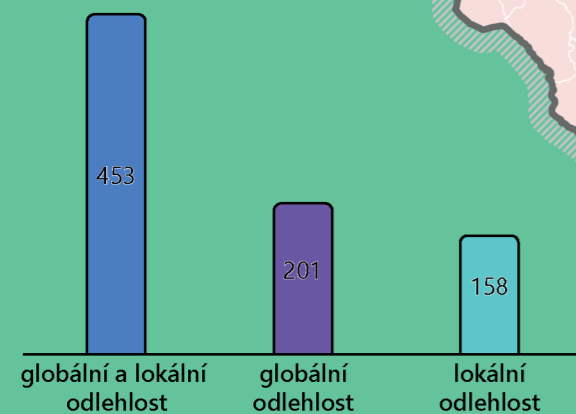
vojenský újezd

## Administrativní členění

- hranice státu
- hranice kraje
- hranice okresu
- hranice krajského města



Počet ZUJ podle typu odlehlosti



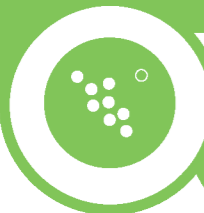
## Příloha 4

Přemysl BONK

Zdroj dat: ArcČR 500 verze 3.3, ČSÚ

Mapové zobrazení: UTM 33N

Katedra geoinformatiky, Olomouc 2024



# ODLEHLÉ HODNOTY PREZIDENTSKÝCH VOLEB

za ZUJ v České republice v roce 2023

Pro výpočet Mahalanobisovi vzdálenosti byla využita funkce **Moutlier**.

Globální Mahalanobisova vzdálenost byla vypočtena z **6 388 prvků**.

Lokální Mahalanobisova vzdálenost byla vypočtena z **220 prvků** zvolených metodou **nejbližšího souseda** každý prvek.

Pro prahovou hodnotu byl použit **95% kvantil Chí kvadrát**.

## Atributy pro výpočet Mahalanobisovi vzdálenosti

Postoupili:	P. Pavel	A. Babiš	
Nepostoupili:	D. Nerudová	P. Fisher	J. Bašta
jiné	volební účast prez. voleb		

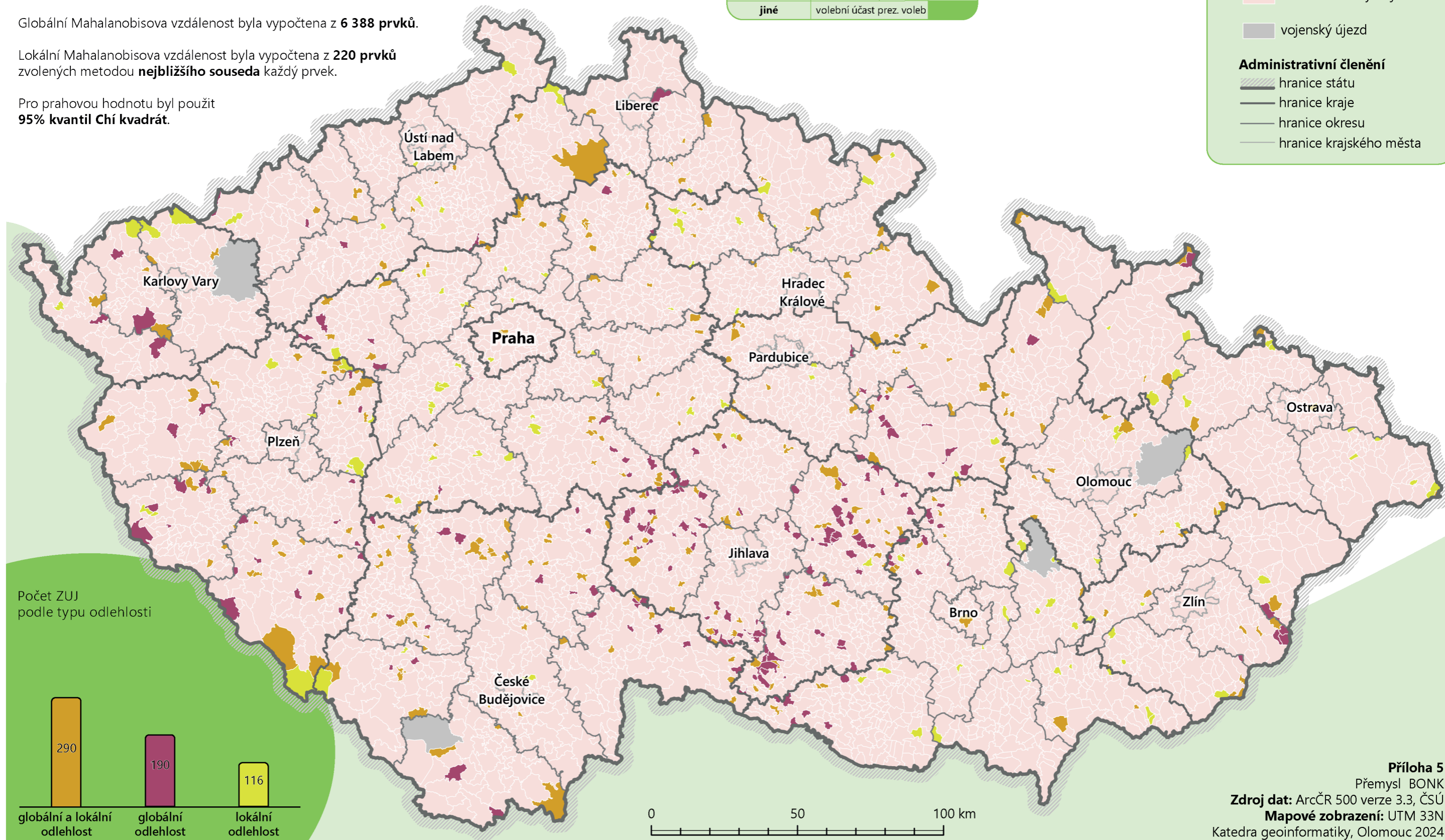
## Klasifikace odlehlosti pro ZUJ

- globální i lokální odlehlost
- globální odlehlost
- lokální odlehlost
- odlehlost nebyla zjištěna

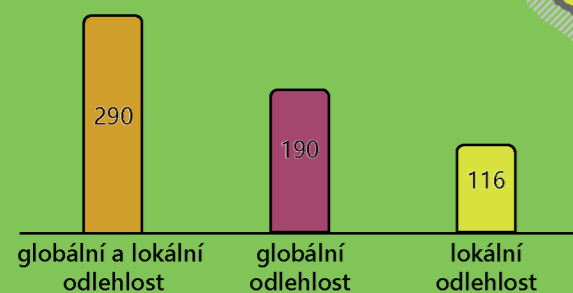
vojenský újezd

## Administrativní členění

- hranice státu
- hranice kraje
- hranice okresu
- hranice krajského města



Počet ZUJ podle typu odlehlosti



## Příloha 5

Přemysl BONK

Zdroj dat: ArcČR 500 verze 3.3, ČSÚ

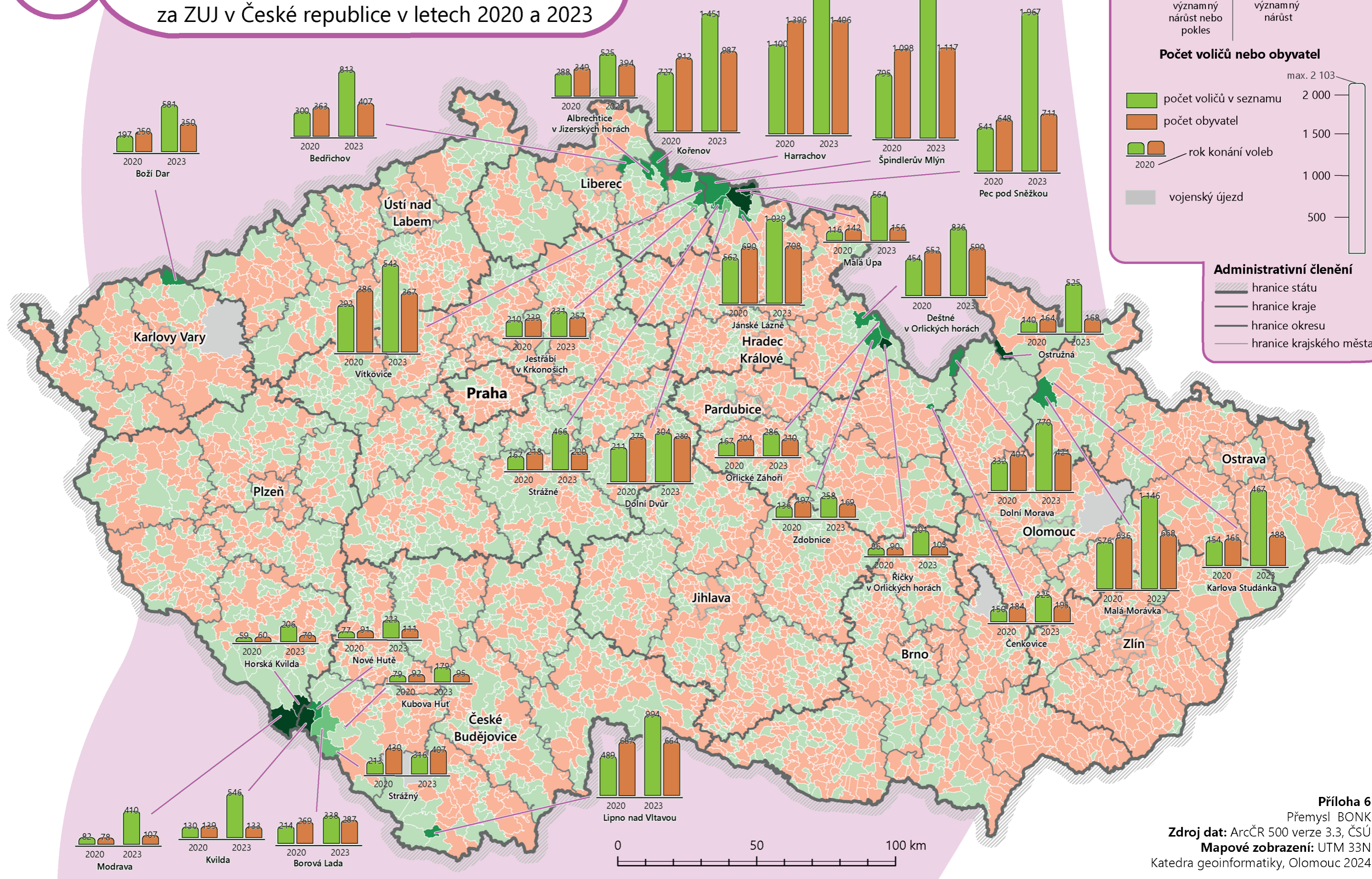
Mapové zobrazení: UTM 33N

Katedra geoinformatiky, Olomouc 2024



# POROVNÁNÍ VOLIČSTVA MEZI VOLBAMI

## za ZUJ v České republice v letech 2020 a 2023







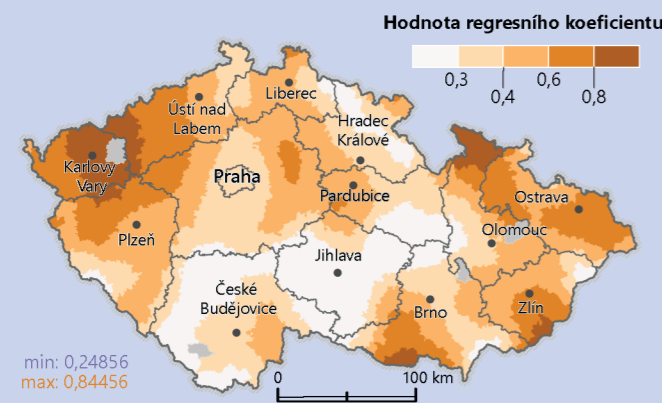
# PROSTOROVÝ REGRESNÍ MODEL KOALICE SPOLU

v České republice v roce 2021

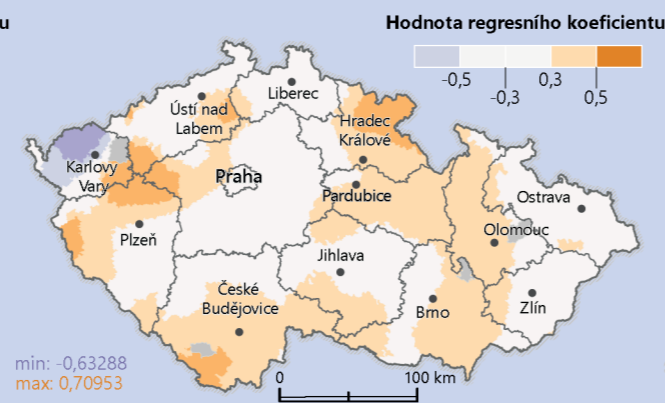
Počet Prediktorů	5
Průměrné skóre R <sup>2</sup> regresního modelu	0,493
Území s mírou lok. kolinearity >30	21,95%

## KOEFICIENTY REGRESNÍHO MODELU

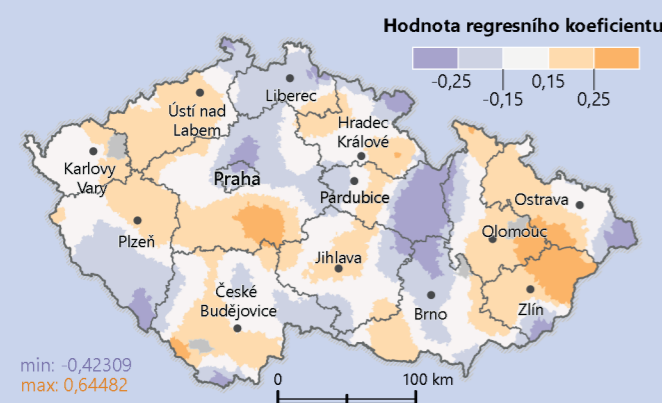
### PODÍL VYSOKOŠKOLSKY VZDĚLANÝCH



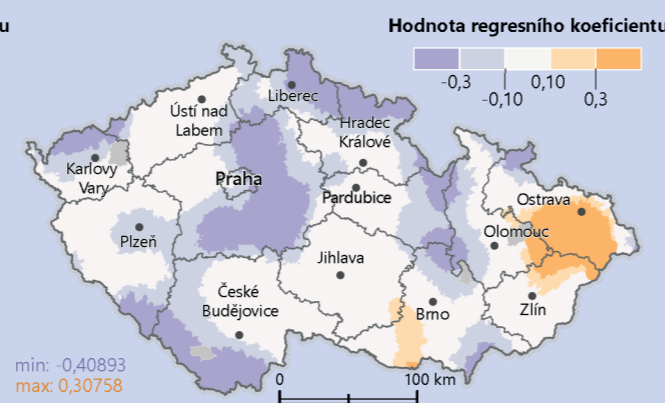
### PODÍL VĚŘÍCÍCH ŘÍMSKOKATOLICKÉ CÍRKVE



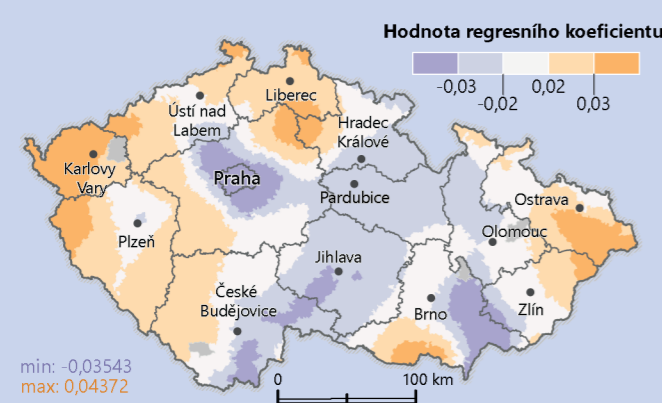
### MÍRA NEZAMĚSTNANOSTI



### PODÍL SEKUNDÁRNÍHO SEKTORU



### INDEX STÁŘÍ

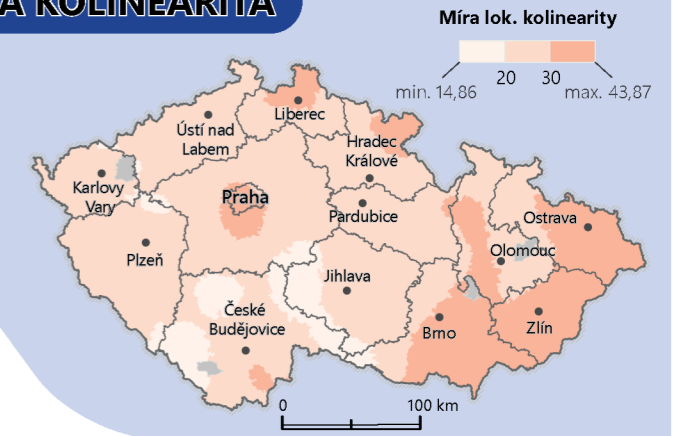


vojenský újezd

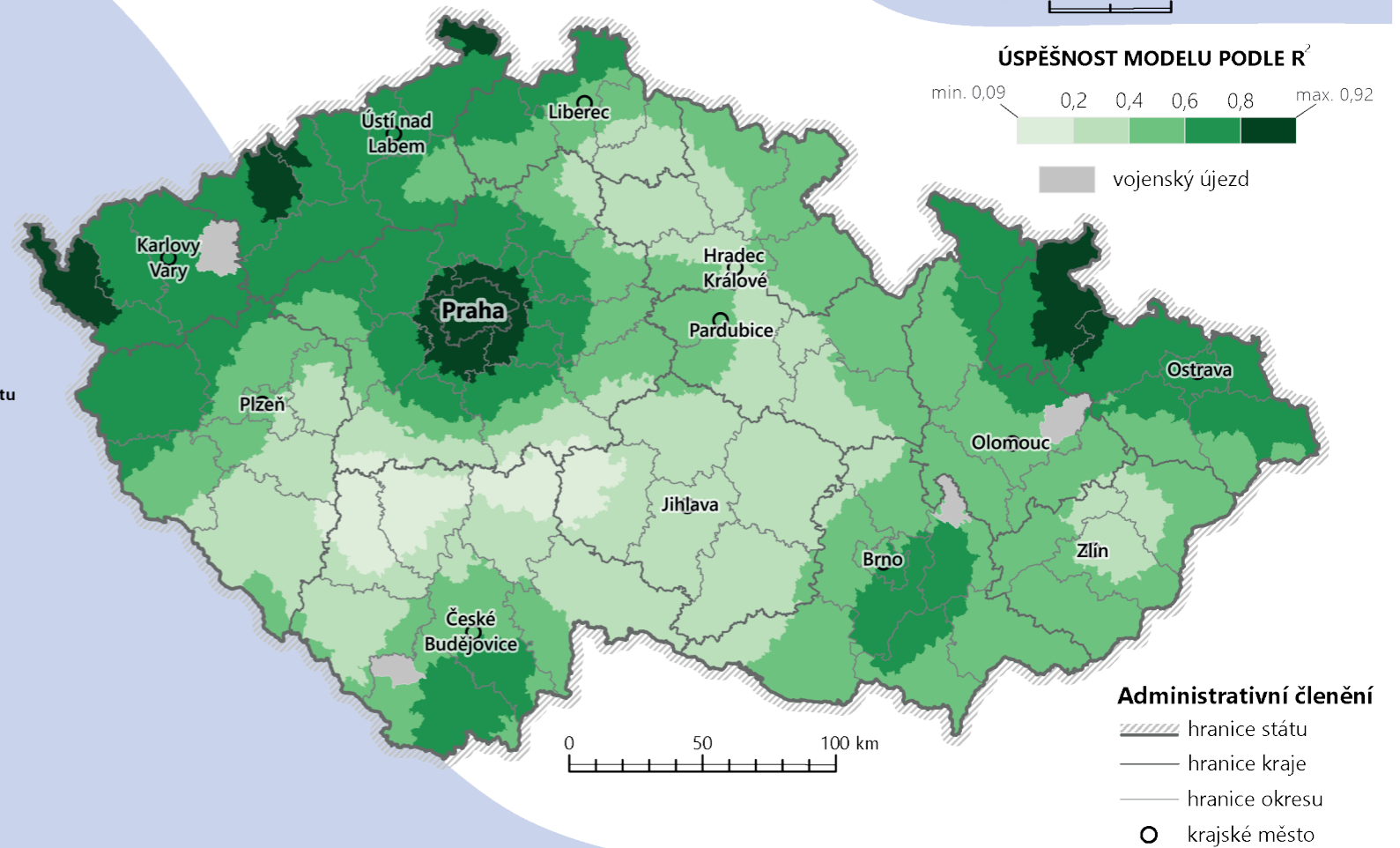
Administrativní členění

hranice státu — hranice kraje • krajské město

## PROSTOROVÁ KOLINEARITA

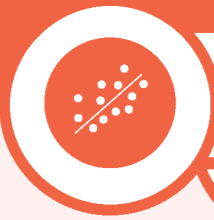


## ÚSPĚŠNOST REGRESNÍHO MODELU



Regresní koeficient vyjadřuje změnu v závislé proměnné v reakci na jednotkovou změnu v příslušného prediktoru, pokud ostatní prediktory zůstanou neměnné.

**Příloha 7**  
Přemysl BONK  
Zdroj dat: ArcČR 500 verze 3.3, ČSÚ, MPSV  
Mapové zobrazení: UTM 33N  
Katedra geoinformatiky, Olomouc 2024



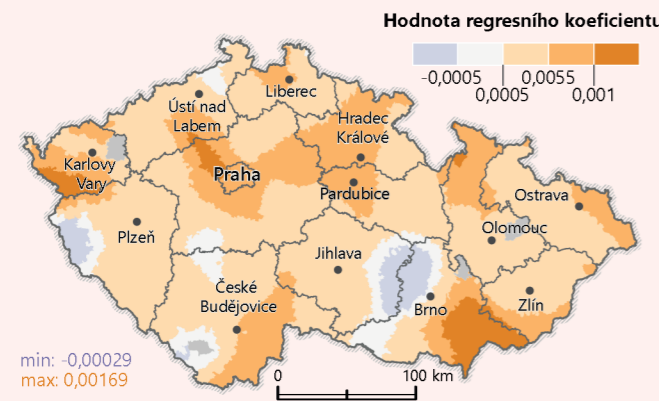
# PROSTOROVÝ REGRESNÍ MODEL Hnutí ANO

v České republice v roce 2021

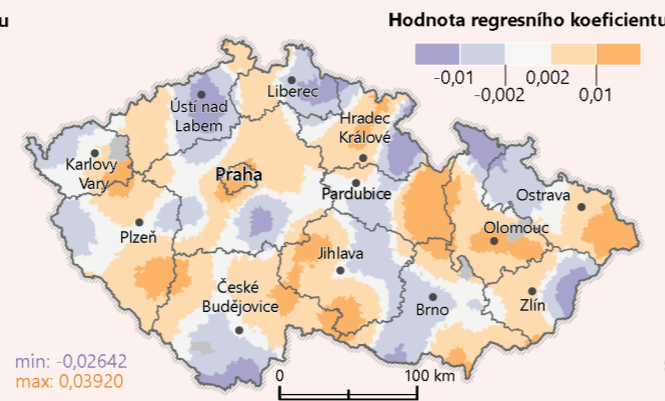
Počet Prediktorů	7
Průměrné skóre R <sup>2</sup> regresního modelu	0,404
Území s mírou lok. kolinearita >30	18,77%

## KOEFICIENTY REGRESNÍHO MODELU

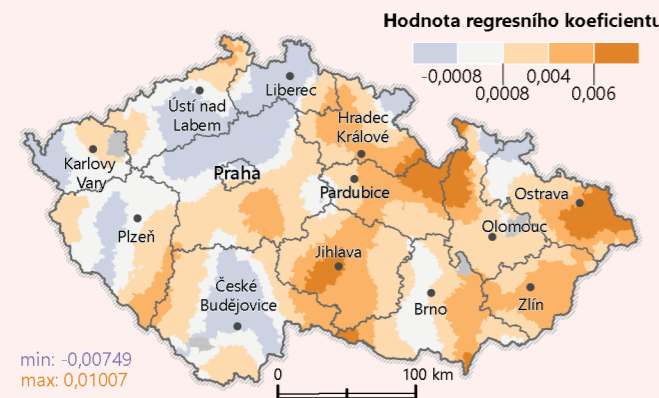
### INDEX STÁŘÍ



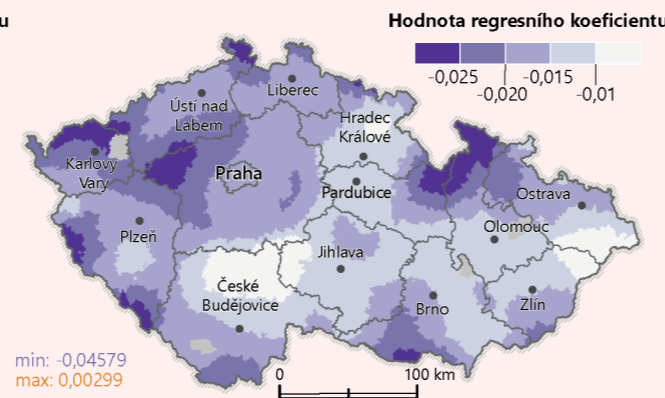
### MÍRA NEZAMĚTNANOSTI



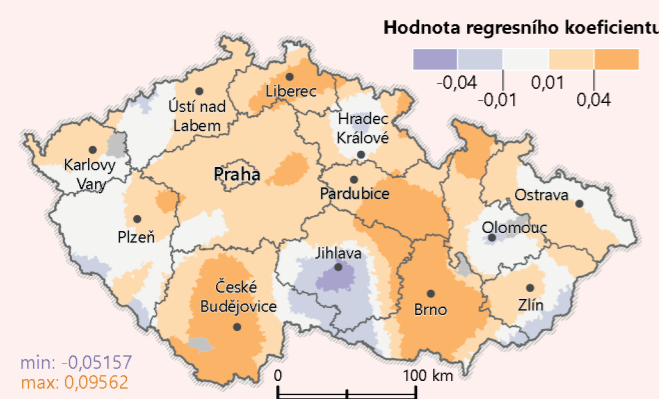
### PODÍL ŽEN SE DVĚMA DĚTI



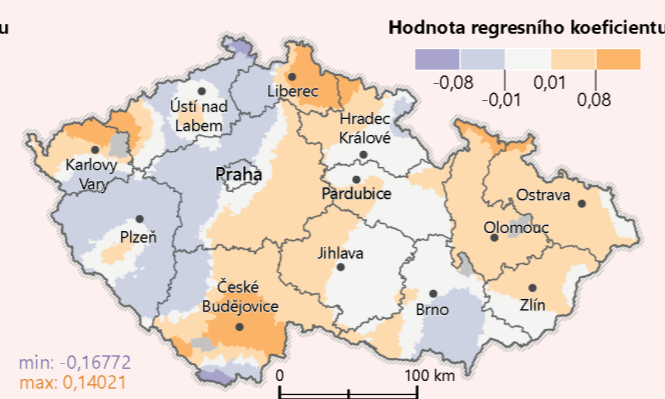
### PODÍL VYSOKOŠKOLSKY VZDĚLANÝCH



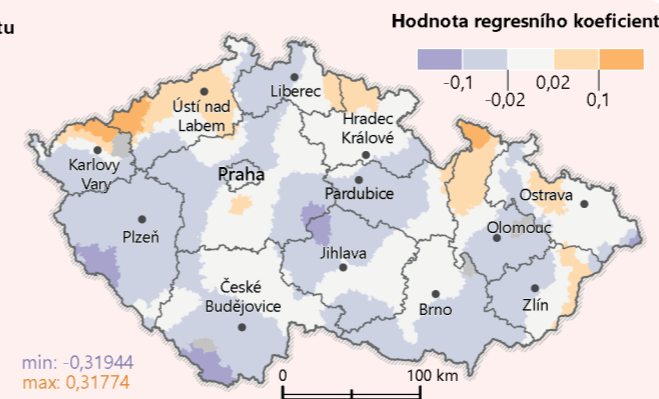
### PODÍL OBYVATEL S PŘÍSPĚVKEM NA BYDLENÍ



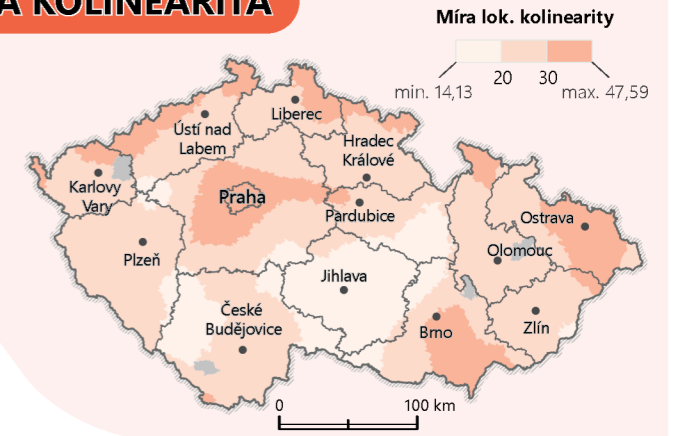
### PODÍL VĚŘÍCÍCH EVANGELICKÉ CÍRKVE



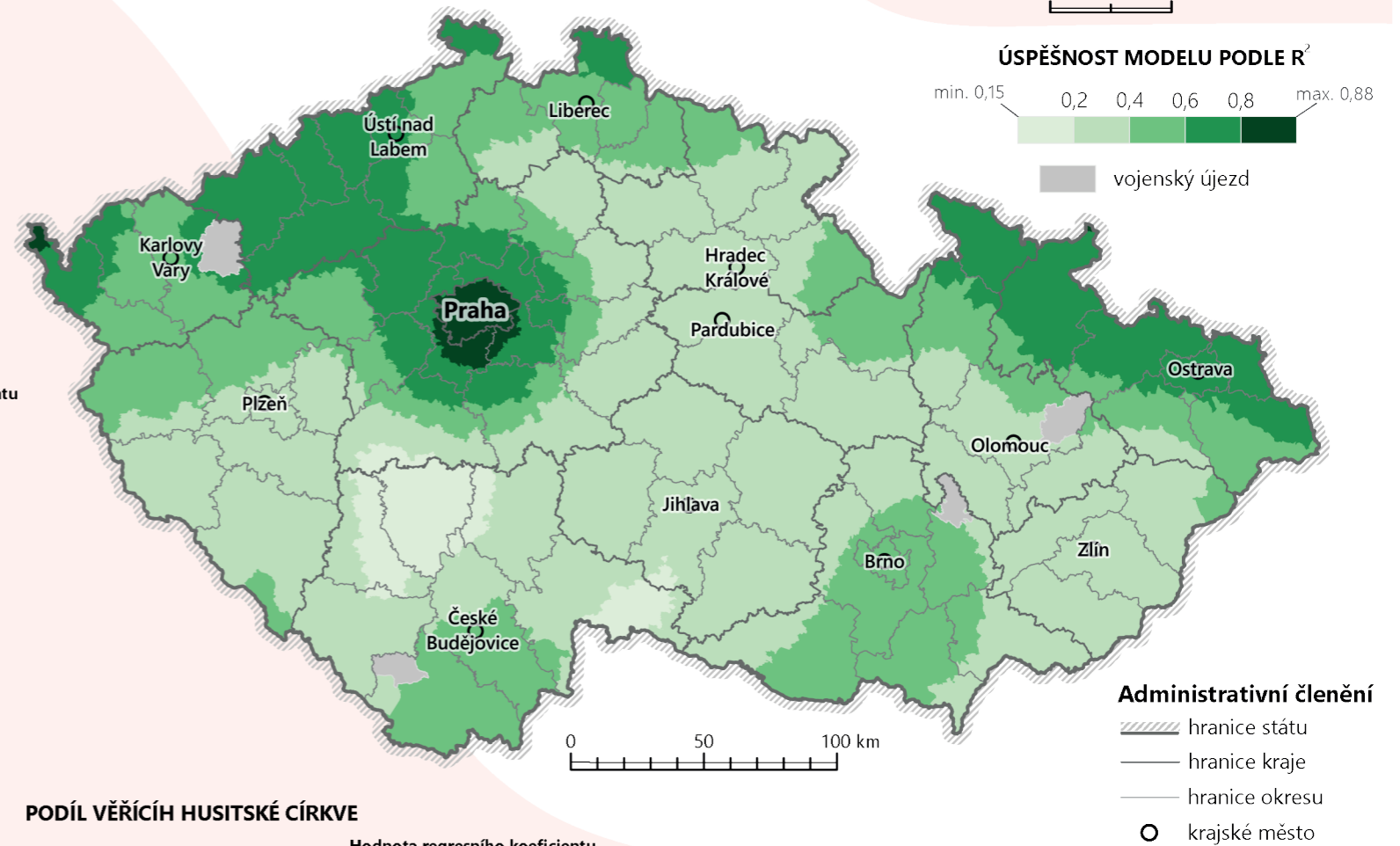
### PODÍL VĚŘÍCÍCH HUSITSKÉ CÍRKVE



## PROSTOROVÁ KOLINEARITA

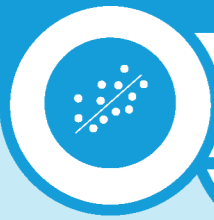


## ÚSPĚŠNOST REGRESNÍHO MODELU



Regresní koeficient vyjadřuje změnu v závislé proměnné v reakci na jednotkovou změnu v příslušného prediktoru, pokud ostatní prediktory zůstanou neměnné.

**Příloha 8**  
Přemysl BONK  
Zdroj dat: ArcČR 500 verze 3.3, ČSÚ, MPSV  
Mapové zobrazení: UTM 33N  
Katedra geoinformatiky, Olomouc 2024



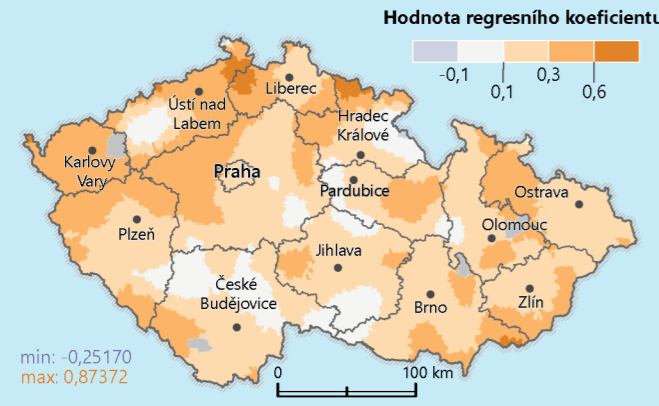
# PROSTOROVÝ REGRESNÍ MODEL KOALICE PIRÁTŮ A STAROSTŮ

v České republice v roce 2021

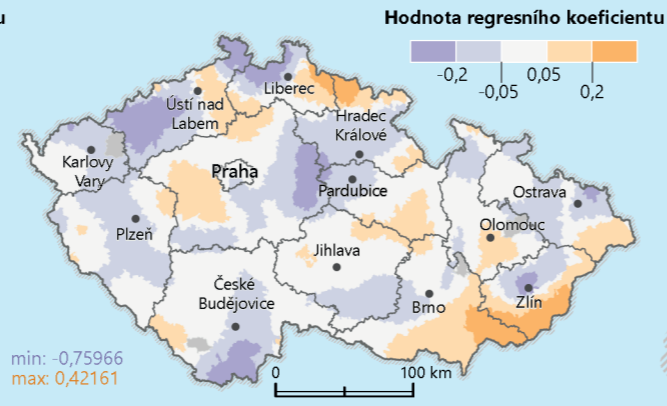
Počet Prediktorů	6
Průměrné skóre R <sup>2</sup> regresního modelu	0,411
Území s mírou kolinearit nad 30°	0,01%

## KOEFICIENTY REGRESNÍHO MODELU

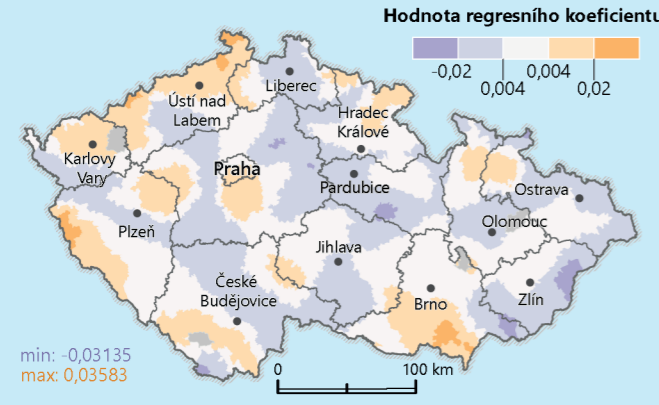
### PODÍL VYSOKOŠKOLSKY VZDĚLANÝCH



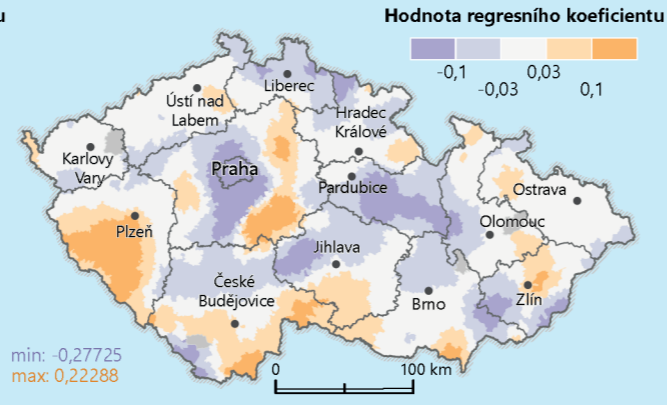
### PODÍL PRIMÁRNÍHO SEKTORU



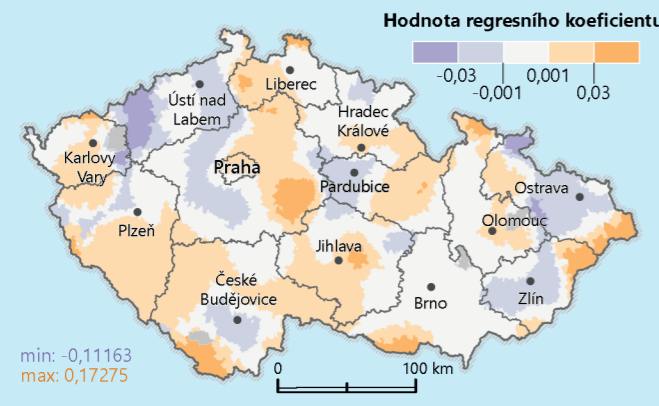
### INDEX STÁŘÍ



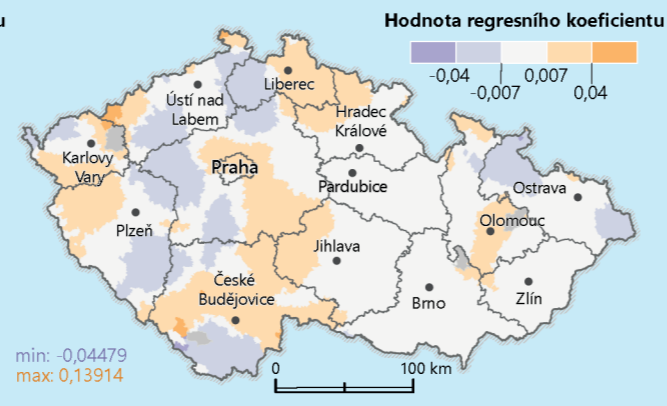
### PODÍL OBYVATEL S PŘÍSPĚVKEM NA BYDLENÍ



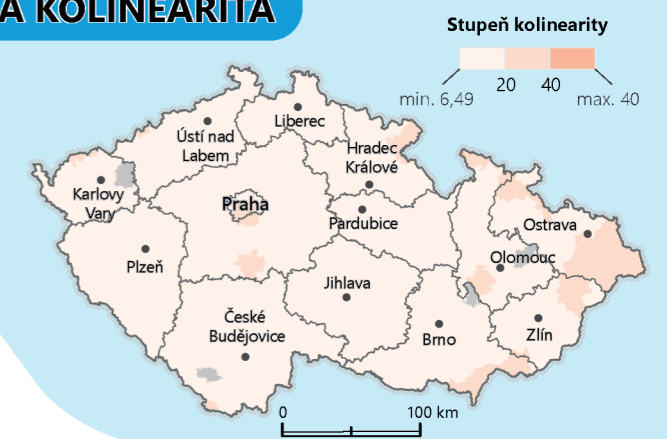
### PODÍL VĚŘÍCÍCH HUSITSKÉ CÍRKVE



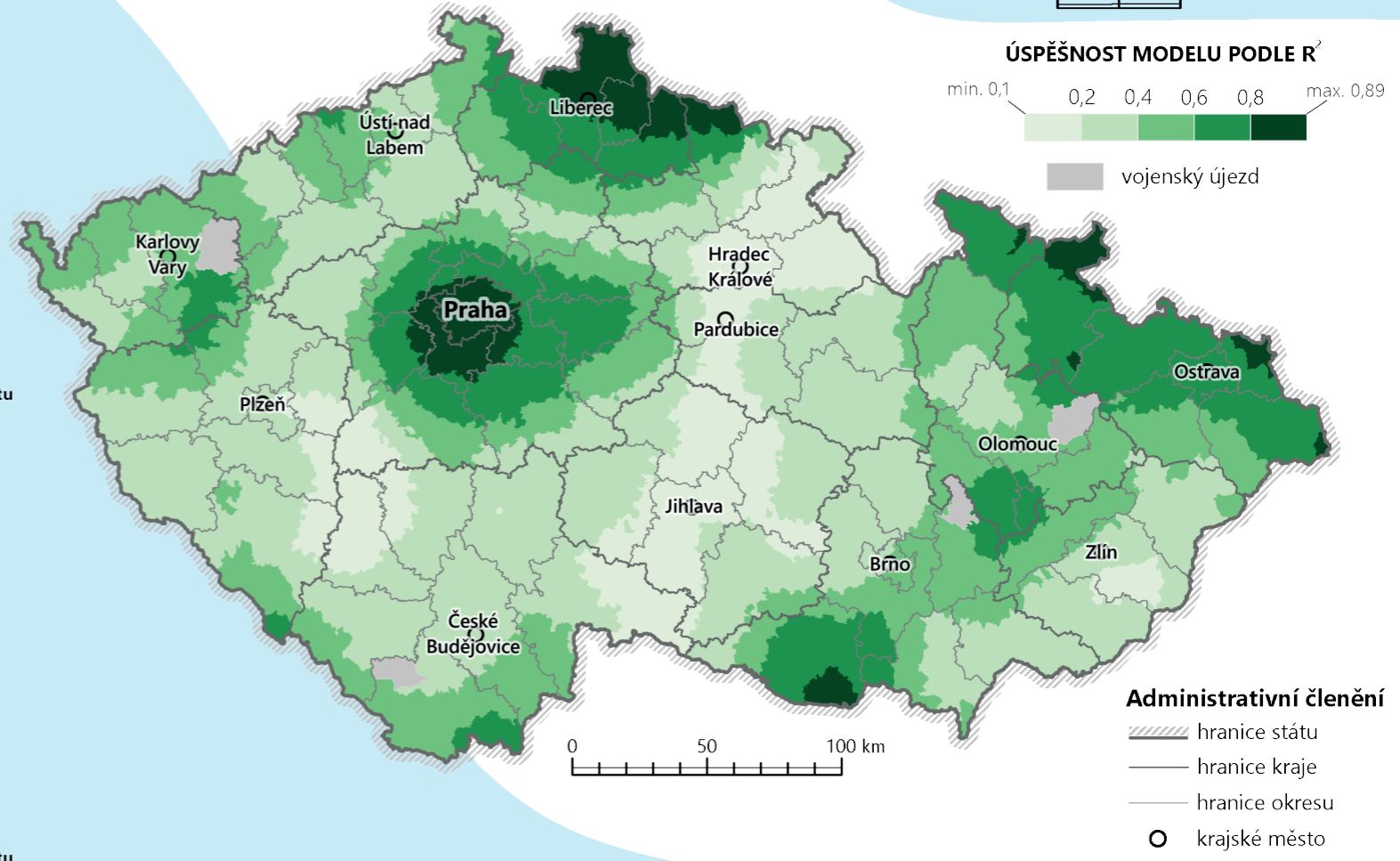
### PODÍL VĚŘÍCÍCH EVANGELICKÉ CÍRKVE



## PROSTOROVÁ KOLINEARITA



## ÚSPĚŠNOST REGRESNÍHO MODELU



vojenský újezd

Administrativní členění

hranice státu — hranice kraje • krajské město

Administrativní členění

hranice státu

hranice kraje

hranice okresu

• krajské město

Regresní koeficient vyjadřuje změnu v závislé proměnné v reakci na jednotkovou změnu v příslušného prediktoru, pokud ostatní prediktory zůstanou neměnné.

**Příloha 9**  
Přemysl BONK  
Zdroj dat: ArcČR 500 verze 3.3, ČSÚ, MPSV  
Mapové zobrazení: UTM 33N  
Katedra geoinformatiky, Olomouc 2024



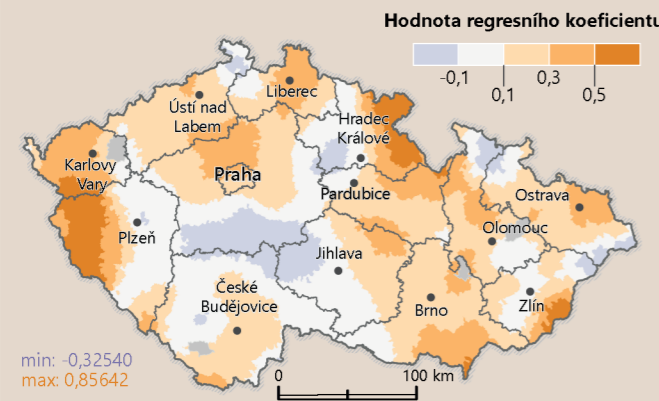
# PROSTOROVÝ REGRESNÍ MODEL Hnutí SPD

v České republice v roce 2021

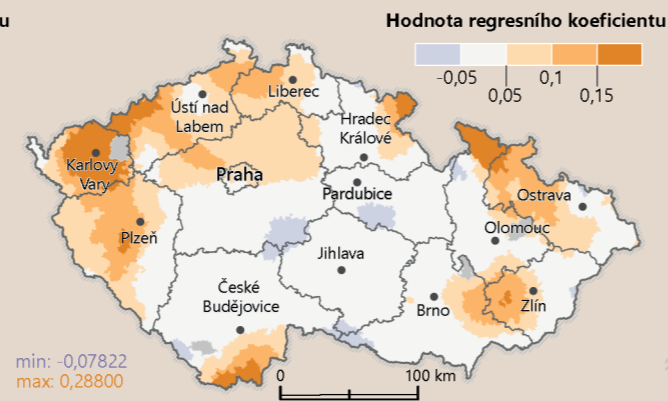
Počet Prediktorů	5
Průměrné skóre R <sup>2</sup> regresního modelu	0,318
Území s mírou kolinearit nad 30°	5,93%

## KOEFICIENTY REGRESNÍHO MODELU

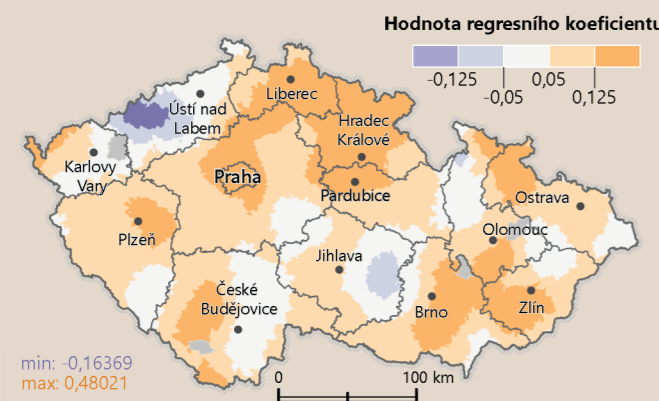
### MÍRA NEZAMĚSTNANOSTI



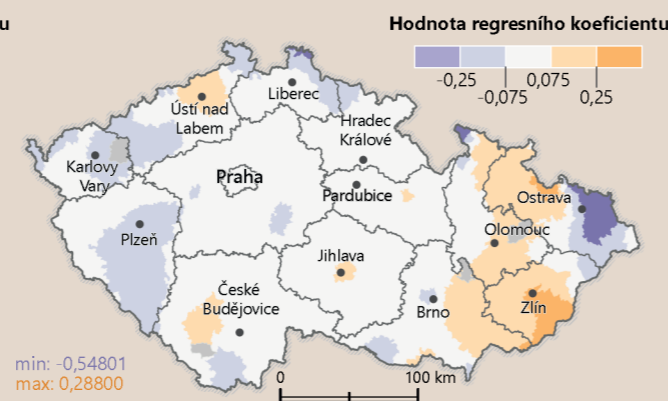
### PODÍL ŽEN SE TŘEMI A VÍCE DĚTMI



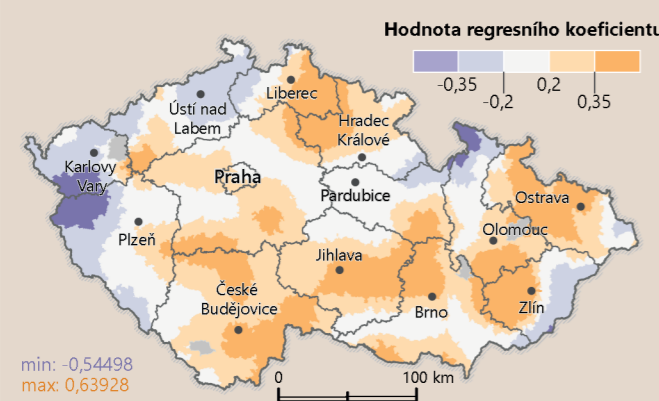
### PODÍL ROZVODŮ



### PODÍL OBYVATEL S OB. BYDLIŠTĚM MIMO KRAJ TB



### PODÍL SEKUNDÁRNÍHO SEKTORU



vojenský újezd

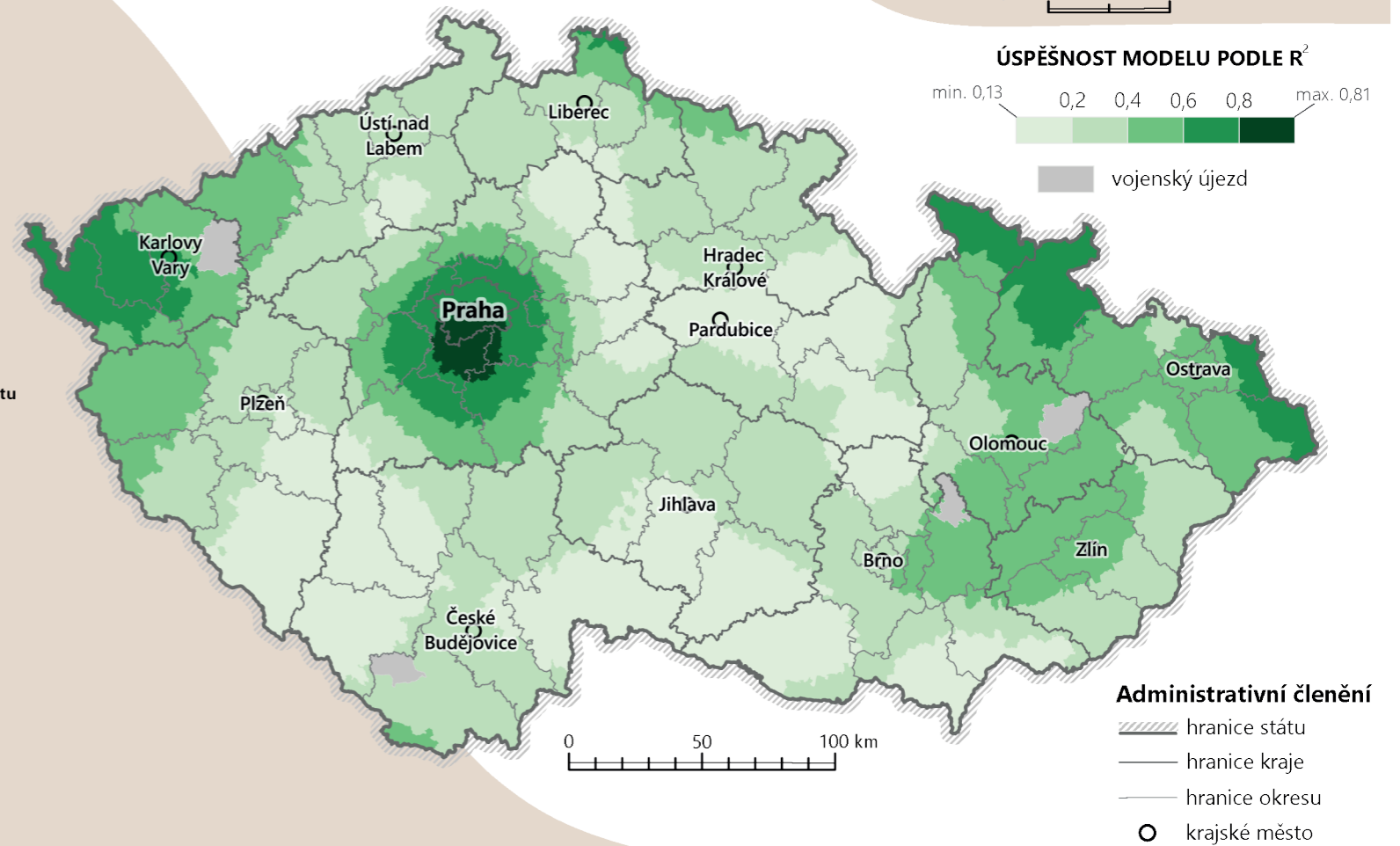
Administrativní členění

hranice státu — hranice kraje • krajské město

## PROSTOROVÁ KOLINEARITA



## ÚSPĚŠNOST REGRESNÍHO MODELU



Regresní koeficient vyjadřuje změnu v závislé proměnné v reakci na jednotkovou změnu v příslušného prediktoru, pokud ostatní prediktory zůstanou neměnné.

**Příloha 10**  
Přemysl BONK  
**Zdroj dat:** ArcČR 500 verze 3.3, ČSÚ, MPSV  
**Mapové zobrazení:** UTM 33N  
Katedra geoinformatiky, Olomouc 2024



# POROVNÁNÍ PREDIKCE MODELŮ LOGISTICKÉ REGRESE

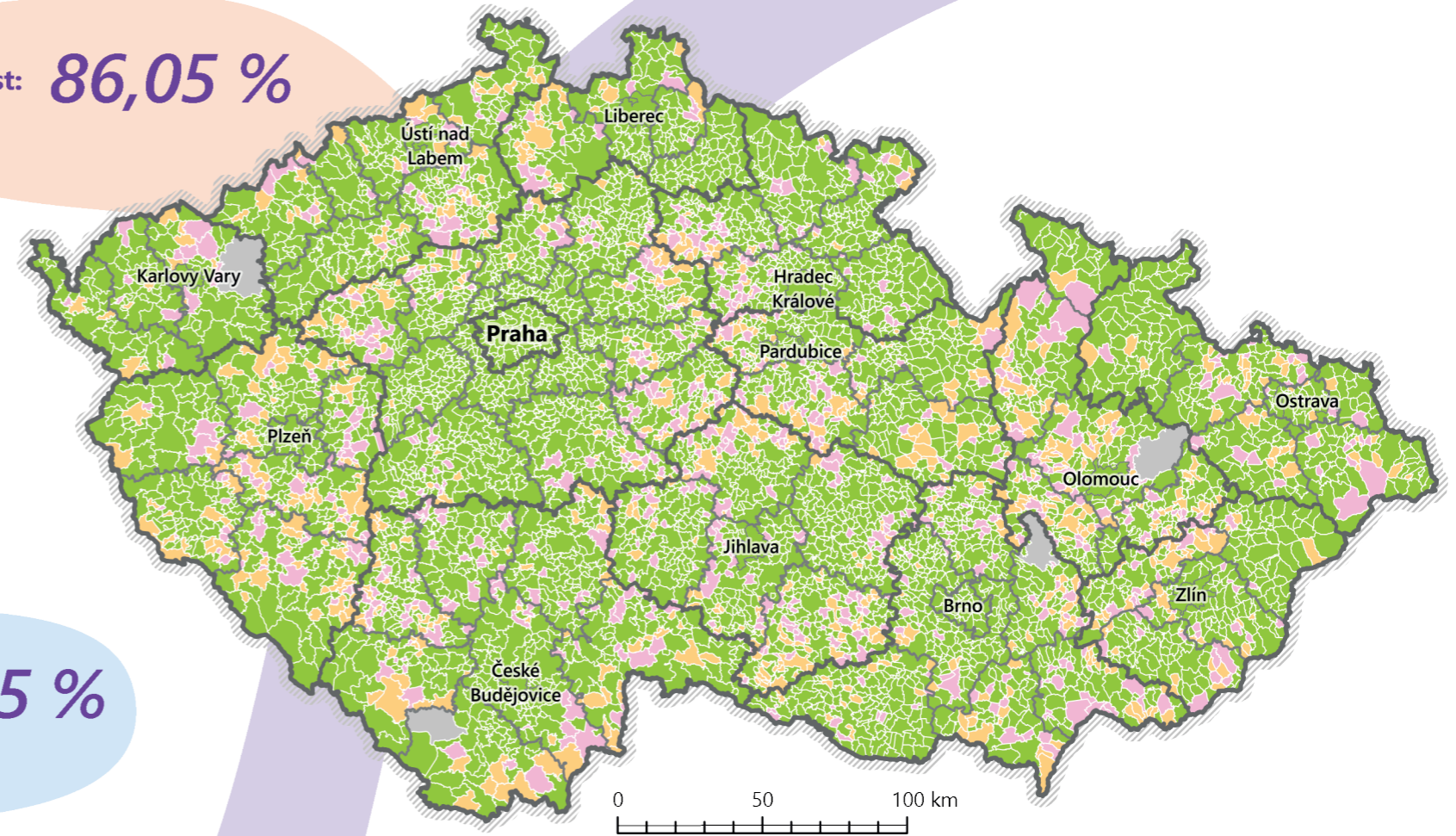
za ZUJ v České republice v roce 2023

PREDIKCE NA ZÁKLADĚ PROSTOROVÝCH LOGISTICKÝCH MODELŮ

MATICE ZÁMĚN Petr Pavel klasický model		Skutečné hodnoty		MATICE ZÁMĚN Andrej Babiš klasický model		Skutečné hodnoty	
		Pozitivní	Negativní			Pozitivní	Negativní
Predikované hodnoty	Pozitivní	2 471 58 %	696 16,3 %	Predikované hodnoty	Pozitivní	675 16 %	362 8,5 %
	Negativní	326 7,8 %	766 17,9 %		Negativní	787 18,5 %	2 435 57 %
MATICE ZÁMĚN Petr Pavel prostorový model		Skutečné hodnoty		MATICE ZÁMĚN Andrej Babiš prostorový model		Skutečné hodnoty	
		Pozitivní	Negativní			Pozitivní	Negativní
Predikované hodnoty	Pozitivní	3 816 59,7 %	729 11,4 %	Predikované hodnoty	Pozitivní	1 385 21,7 %	383 6 %
	Negativní	393 6,2 %	1 450 22,7 %		Negativní	794 12,4 %	3 826 59,9 %

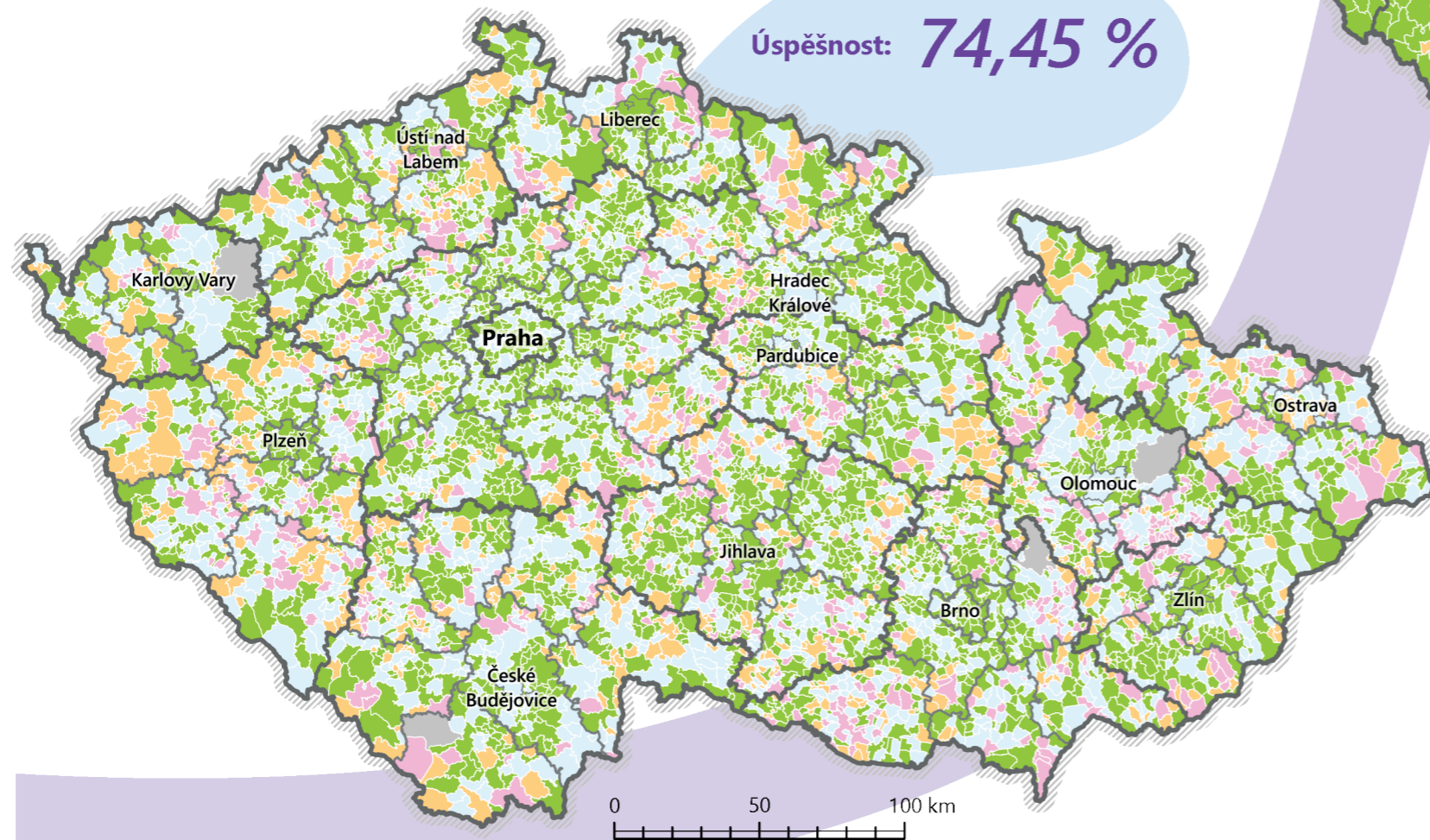
Úspěšnost: **86,05 %**

Vstupní prediktory logistických modelů	
Petr Pavel	Andrej Babiš
podíl zaměstnaných osob	index stáří
index kvality života	podíl osob se středoškolským vzděláním bez maturity
vysokoškolské vzdělání	podíl nevěřících obyvatel
podíl obyvatel v exekuci	podíl obyvatel s příspěvkem na bydlení
podíl osob s OB v jiném kraji než TB	podíl svobodných osob



PREDIKCE NA ZÁKLADĚ KLASICKÝCH LOGISTICKÝCH MODELŮ

Úspěšnost: **74,45 %**



### Úspěšnost predikce podle logistických modelů za ZUJ

- správná predikce v obou případech
- správná predikce jen v 1 případě
- špatná predikce v obou případech

■ trénovací data

■ vojenský újezd

### Administrativní členění

- hranice státu
- hranice kraje
- hranice okresu
- hranice krajského města

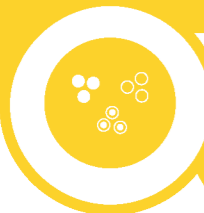
Příloha 11

Přemysl BONK

Zdroj dat: ArcČR 500 verze 3.3, ČSÚ, MPSV

Mapové zobrazení: UTM 33N

Katedra geoinformatiky, Olomouc 2024



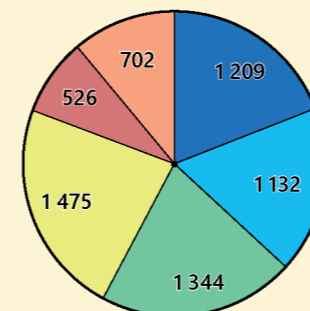
# CHARAKTERISTIKA A PREFERENCE SHLUKŮ

za ZUJ v České republice v roce 2021

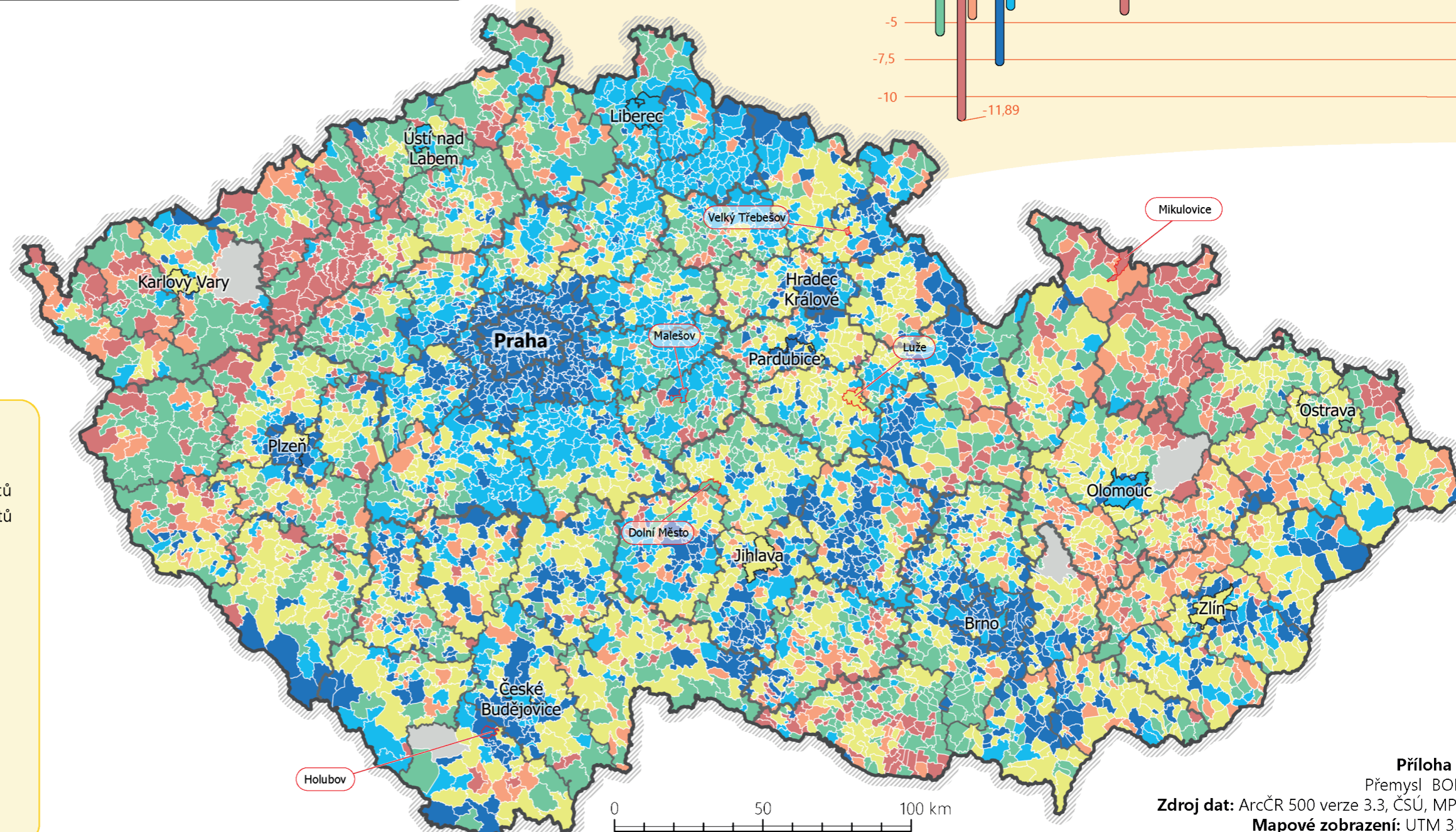
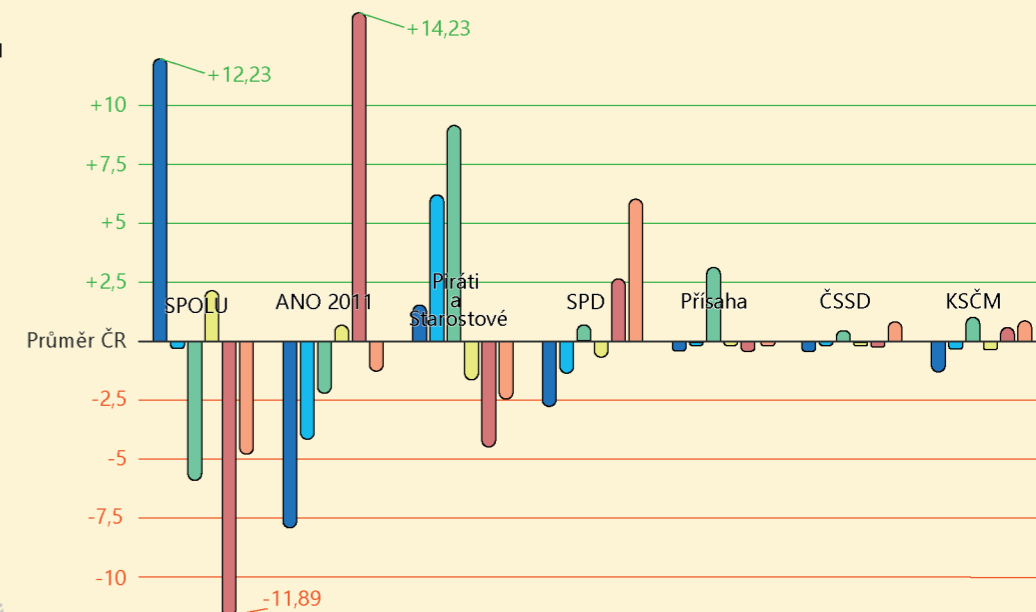
Hodnoty vybraných SED ukazatelů shluků vůči celostátnímu průměru

Typická ZUJ	podpora vládní koalice	podpora koalice Pirátů a starostů	podpora koalice Pirátů a Starostů a hnutí Přísaha	volební zisk odpovídající průměrným hodnotám ČR	podpora vládní opozice	podpora hnutí SPD
Holubov	Malešov	Dolní Město	Luže	Mikulovice	Velký Třebešov	
index stáří	▼ nižší	— střední	▲ vyšší	— střední	▲ vyšší	— střední
index kvality života	▲ vyšší	— střední	▼ nižší	— střední	▼ nižší	— střední
index soc. vyloučení	▼ nižší	— střední	▼ nižší	— střední	▲ vyšší	▲ vyšší
index města a venkova	▲ vyšší	▼ nižší	— střední	— střední	▼ nižší	— střední
počet obyvatel	▲ vyšší	▼ nižší	— střední	— střední	▼ nižší	— střední

Počet ZUJ zařazených do shluku



Rozdíly průměrů volebních zisků mezi shluky a celou ČR v %



## Charakteristika volebního chování ZUJ podle shlukové analýzy

- podpora vládní koalice
- podpora koalice Pirátů a Starostů
- podpora koalice Pirátů a Starostů a hnutí Přísaha
- volební zisk odpovídající průměrným hodnotám ČR
- podpora vládní opozice
- podpora hnutí SPD
- hranice ZUJ typické pro shluk
- vojenský újezd

## Administrativní členění

- hranice státu
- hranice kraje
- hranice okresu
- hranice krajského města