

PROSTOROVÉ HODNOCENÍ DAT VOLEBNÍCH VÝSLEDKŮ

CÍLE PRÁCE

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

V první části práce byla sestavena datová sada volně dostupných dat. Primárním cílem bylo shromáždění dat výsledků voleb do poslanecké sněmovny a voleb prezidenta republiky. Vytvořená datová sada byla dále doplněna o vhodné sociální ekonomické a demografické ukazatele.

V druhé části diplomové práce byly aplikovány vhodné prostorové statistické metody za účelem hledání odlehých hodnot nebo identifikace shluků podobných hodnot, nebo sestavení modelů, které budou popisovat vztahy mezi volebním výsledkem a SED situací. Hlavní důraz byl při zpracování kladen především na využití prostorových statistických metod.

Ve třetí části diplomové práce budou výsledky analýz vhodně vizualizovány a výsledky provedených analýz konzultovány s politologem.

POSTUP ZPRACOVÁNÍ

1. PŘÍPRAVNÉ PRÁCE

studium literatury

sestavení datové sady

úprava vstupních dat

exploratorní analýza

2. APLIKACE ANALÝZ

analýza Mahalanobisovi vzdálenosti

analýza voličských průkazů

geograficky vážená regrese

logistická regrese

shluková analýza

3. HODNOCENÍ, VIZUALIZACE A INTERPRETACE

vyhodnocení analýz

vizualizace výsledků

interpretace výsledků

4. ZÁVĚREČNÉ PRÁCE

dokončení textu diplomové práce

vytvoření posteru

vytvoření webových stránek

POUŽITÉ METODY

Mahalanobisova vzdálenost

Mahalanobisova vzdálenost je používána pro vyhledání odlehých hodnot v rozsahu definovaných atributů datové sady. Při analýze volebních výsledků byly s pomocí Mahalanobisovy vzdálenosti odhaleny odlehle hodnoty dat volebních výsledků.

Lineární regrese a GWR

Lineární regrese je metoda, jejíž principem je vysvětlení vztahu mezi závislou proměnnou s pomocí vysvětlujících proměnných. 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.

Geograficky vážená regrese je typ regresního modelu, který popisuje vztah mezi závislou proměnnou a vysvětlujícími proměnnými v závislosti na geografickém umístění. V diplomové práci 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.

Logistická regrese

Logistická regrese je statistickou metodou, která je využívána k modelování hodnot závislé proměnné binárního charakteru. V analýze volebních výsledků byla porovnána míra úspěšnosti predikce na základě klasických a prostorově vážených modelů logistické regrese.

Shluková analýza

Shluková analýza je statistická metoda, kterou jsou na základě určité metody či algoritmu prvky zařazeny do shluků společně s prvky podobných vlastností. V diplomové práci byla tato metoda použita k vymezení shluků podle volebního chování, jež byly následně charakterizovány dalšími SED ukazateli.

VÝSLEDKY PROVEDENÝCH ANALÝZ

Analýza odlehých hodnot

V první provedené analýze byla hodnocena mnohorozměrná míra odlehlosti s využitím Mahalanobisovi vzdálenosti z globálního i lokálního pohledu u sněmovních voleb 2021 i prezidentských voleb 2023. Při hodnocení bylo zjištěno, že **kraj Vysočina** obsahuje **nejvíce ZUJ klasifikovaných jako odlehle**. Dalším faktem je **vyšší výskyt odlehých hodnot** v příhraničních oblastech a oblastech **vnitřních periferií na hranicích krajů**.

Analýza využití voličských průkazů

V rámci druhé analýzy bylo zkoumáno využití voličských průkazů v kontextu prezidentských voleb 2023. S využitím informací o voličích v seznamu, na kterých je založený index 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.

Vysvětlení volebních výsledků s pomocí GWR modelů

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. 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**. Dále jsou v modelech zaznamenány informace o lokální kolinearitě a úspěšnosti modelu ve vysvětlení prostorové variability závislé proměnné. Nej přesnějším GWR modelem z hlediska vysvětlení variability byl GWR model **koalice SPOLU**. Modelem s nejnižší lokální kolinearitou byl GWR model **koalice Pirátů a Starostů**.

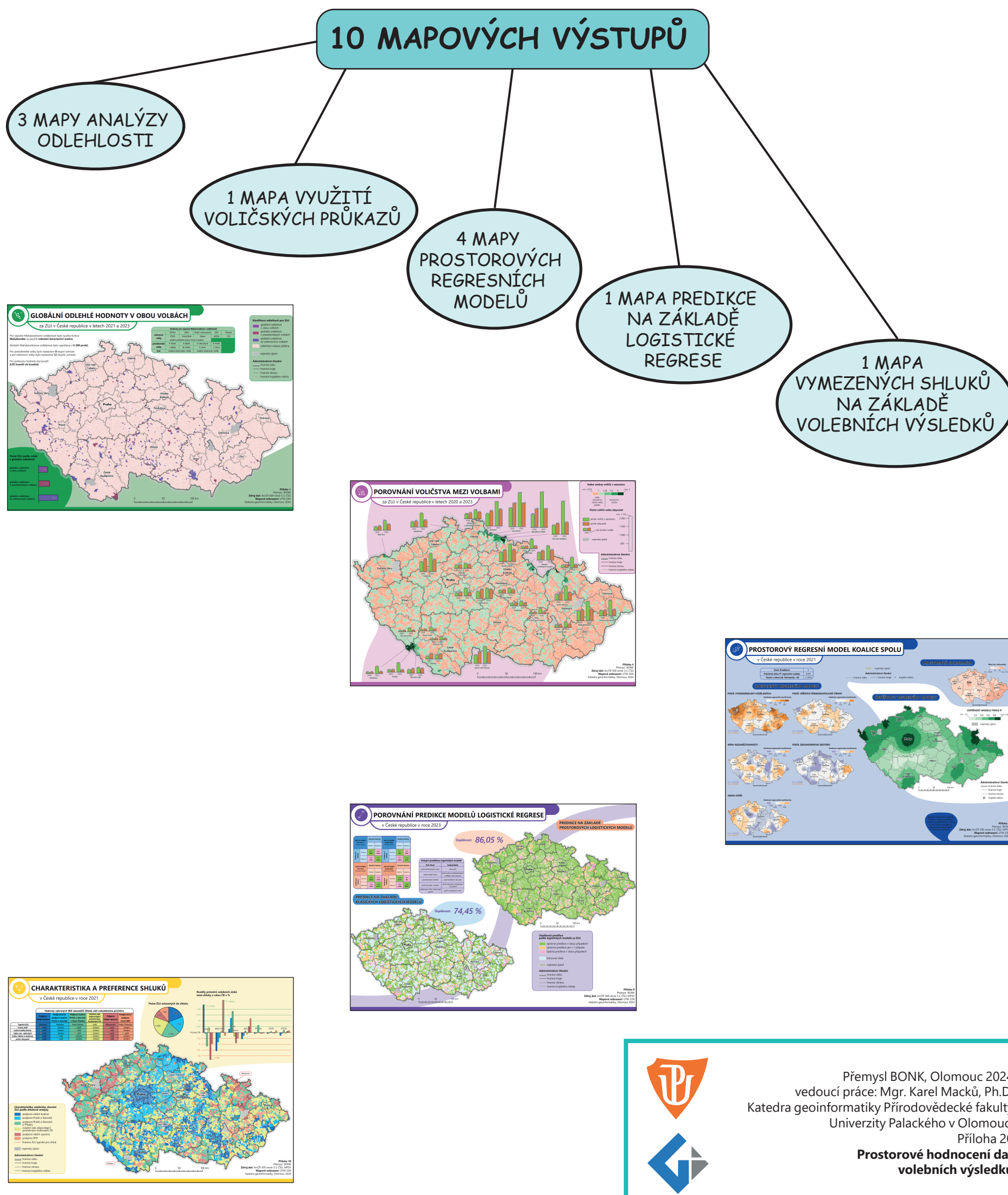
Analýza přesnosti predikce na základě modelů logistické regrese

Čtvrtá zpracovaná analýza se zabývala porovnáním úspěšnosti predikcí v druhém kole prezidentských voleb 2023, založených na klasických a prostorově vážených modelech logistické regrese. Z hodnocení přesnosti provedených predikcí 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.

Shluková analýza a následné vymezení volebního chování

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. 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 celostátního průměru. Výsledné rozdíly SED ukazatelů v dalším kroku rozděleny na nadprůměrné, střední a nadprůměrné hodnoty.

VÝSTUPY DIPLOMOVÉ PRÁCE



Přemysl BONK, Olomouc 2024
vedoucí práce: Mgr. Karel Macků, Ph.D.
Katedra geoinformatiky Přírodovědecké fakulty
Univerzity Palackého v Olomouci
Příloha 20
Prostorové hodnocení dat volebních výsledků