



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**

FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ**

DEPARTMENT OF COMPUTER SYSTEMS

**ANALÝZA TEPLOTNÍ MAPY MĚSTA BRNA**

DATA ANALYSIS OF HEATMAP OF BRNO

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

**JAKUB KOZUBEK**

**VEDOUCÍ PRÁCE**

SUPERVISOR

**Ing. VOJTĚCH MRÁZEK, Ph.D.**

BRNO 2023

## Zadání bakalářské práce



142348

Ústav: Ústav počítačových systémů (UPSY)  
Student: **Kozubek Jakub**  
Program: Informační technologie  
Specializace: Informační technologie  
Název: **Analýza teplotní mapy města Brna**  
Kategorie: Data mining  
Akademický rok: 2022/23

### Zadání:

1. Seznamte se s datovými zdroji pro město Brno, zejména s teplotními mapami.
2. Vyhledejte další datové zdroje, které by mohly vést k vysvětlení ohřívání jednotlivých částí města.
3. Zpracujte studii na výše uvedené téma.
4. Stanovte hypotézy způsobující ohřívání jednotlivých míst města.
5. Hypotézy experimentálně ověřte.
6. Vyhodnoťte závěry ze získaných hypotéz.

### Literatura:

- Dle pokynů vedoucího.

Při obhajobě semestrální části projektu je požadováno:

- Splnění bodů 1 až 3 zadání.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Mrázek Vojtěch, Ing., Ph.D.**  
Vedoucí ústavu: Sekanina Lukáš, prof. Ing., Ph.D.  
Datum zadání: 1.11.2022  
Termín pro odevzdání: 10.5.2023  
Datum schválení: 31.10.2022

## Abstrakt

Práce se zabývá analýzou teplotních dat pro oblast města Brna. Věnuje se návrhu a ověřování hypotéz pro oteplování a ochlazování částí města. Výsledkem práce jsou metody pro statistické testování a pro regresní analýzu implementované v programovacím jazyce Python, výsledky získané pomocí těchto metod a jejich následná interpretace vzhledem k původně navrženým hypotézám.

## Abstract

The thesis deals with the analysis of temperature data for the area of the city of Brno. It is focused on the design and verification of hypotheses for warming and cooling parts of the city. The result of the work are methods for statistical testing and for regression analysis implemented in the Python programming language, the results obtained using these methods and their subsequent interpretation in relation to the originally proposed hypotheses.

## Klíčová slova

geografická data, teplotní analýza, městský tepelný ostrov, regrese, prostorová regrese, statistické testy, Python, PySAL, Brno

## Keywords

geographic data, temperature analysis, urban heat island, regression, spatial regression, statistical tests, Python, PySAL, Brno

## Citace

KOZUBEK, Jakub. *Analýza teplotní mapy města Brna*. Brno, 2023. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Vojtěch Mrázek, Ph.D.

# Analýza teplotní mapy města Brna

## Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením pana doktora Mrázka. Uvedl jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpal.

.....  
Jakub Kozubek  
4. května 2023

## Poděkování

Rád bych poděkoval panu Ing. Vojtěchu Mrázkovi, Ph.D. za vedení této práce, užitečné rady a rychlou zpětnou vazbu. Taktéž bych chtěl poděkovat svojí partnerce, spolužákům a rodině za podporu a cenné rady.

# Obsah

<b>1</b>	<b>Úvod</b>	<b>3</b>
<b>2</b>	<b>Geografická data a jejich formát</b>	<b>5</b>
2.1	Geografická data . . . . .	5
2.2	Kartografické zobrazení a souřadnicový systém . . . . .	5
2.3	GIS software . . . . .	7
2.4	Formáty pro ukládání dat . . . . .	8
2.4.1	GeoJSON . . . . .	8
2.4.2	Shapefile . . . . .	9
2.4.3	Formáty založené na XML . . . . .	9
<b>3</b>	<b>Metody pro ověřování stanovených hypotéz</b>	<b>11</b>
3.1	Statistické testování hypotéz . . . . .	11
3.2	Statistické testy . . . . .	12
3.3	Regrese a prostorová regrese . . . . .	13
<b>4</b>	<b>Návrhy hypotéz a způsobu jejich ověření</b>	<b>16</b>
4.1	Způsob tvorby hypotéz . . . . .	16
4.2	Dostupné datové sady . . . . .	16
4.2.1	Teplota povrchu . . . . .	17
4.2.2	Industriální areály, brownfieldy a skládky . . . . .	18
4.2.3	Využití půdy . . . . .	19
4.2.4	Druh městské zástavby . . . . .	21
4.2.5	Významné plochy zeleně . . . . .	22
4.2.6	Intenzita dopravy . . . . .	22
4.2.7	Zelené střechy . . . . .	23
<b>5</b>	<b>Implementace ověřování hypotéz a jejich testování</b>	<b>24</b>
5.1	Postup při využití statistických testů a porovnání průměrů . . . . .	24
5.2	Výsledky testování hypotéz pomocí statistických testů . . . . .	26
5.2.1	Městská zástavba . . . . .	26
5.2.2	Využití půdy . . . . .	29
5.2.3	Industriální areály, skládky a brownfieldy . . . . .	30
5.2.4	Různé další datové sady . . . . .	32
5.3	Postup při využití regrese a prostorové regrese . . . . .	34
5.3.1	Výsledky regresní analýzy . . . . .	35
5.4	Vybrané potvrzené teorie podrobněji . . . . .	40
5.4.1	Lomy . . . . .	40

5.4.2	Významné zelené plochy . . . . .	41
5.4.3	Industriální areály . . . . .	42
<b>6</b>	<b>Závěr</b>	<b>44</b>
	<b>Literatura</b>	<b>45</b>

# Kapitola 1

## Úvod

Teplota a oteplování je téma, které lze nalézt téměř na každém rohu, v médiích nebo například i když si během vlny veder v létě uvědomíme, jaké horko nám je a jak na nás působí. Ve městské zástavbě je tento efekt ještě výraznější, je to jeden z nejvíce zdokumentovaných jevů klimatické změny. Jedná se o zvýšenou teplotu v městských oblastech oproti oblastem mimo město - tzv. teplotní ostrov [21]. Porozumět tomuto jevu a co má na teplotu největší vliv může být velmi důležité pro územní plánování, lidské zdraví nebo životní prostředí. Tato práce se zaměřuje na oblast města Brna s využitím volně přístupných dat poskytovaných magistrátem. Identifikace vzorců a faktorů ovlivňujících rozložení teploty ve městech jim může pomoci se adaptovat na tyto výzvy způsobené klimatickou změnou jako jsou v posledních letech často delší a intenzivnější vlny veder. Může také pomoci s rozvojem udržitelného městského designu, jako je zelená infrastruktura nebo například energeticky účinné budovy, které povedou ke zmírnění problémů způsobených efektem teplotního ostrova.

Několik studií zabývajících se efektem teplotního ostrova bylo provedeno na případech různých světových měst. Například [17] se zabývá fyzikální podstatou tohoto efektu a různými fyzikálními veličinami a povětrnostními vlivy, které k efektu teplotního ostrova přispívají. [24] zkoumá příčiny způsobující oteplování v zastavěných oblastech, konkrétně ve 30 velkých čínských městech. V evropském prostředí se tímto tématem zabýval [22], a to konkrétně možným snižováním dopadů tohoto efektu nebo způsoby jakým mu předcházet a snižovat jeho přítomnost.

První dvě výše zmíněné studie se zabývají širokým pohledem na tuto problematiku, kdy zkoumají např. velikost města, jeho tvar a celkové rozložení. Třetí studie se zabývá konkrétními prvky, které by mohly mít schopnost tento efekt ovlivňovat. Tato práce se snaží přinést podrobnější pohled na tuto problematiku. Bude zkoumat vliv konkrétních prvků – vrstev na efekt teplotního ostrova, a to jak vrstev, které mají pozitivní, tak negativní vliv na teplotu. Pro toto zkoumání bude využito statistických testů k porovnání teplot a také regresní analýzy ke zjištění míry vlivu jednotlivých činitelů, která byla již podobným způsobem využita v [13].

Výsledky této práce mohou přispět k porozumění, jaké konkrétní prvky mohou ve městě Brně, případně i v dalších městech s podobnými klimatickými podmínkami, oteplovat, případně ochlazovat své okolí. Tyto poznatky mohou být využity při výstavbách nových částí nebo přestavbách těch stávajících. Navržené metody lze teoreticky využít pro analýzu dalších měst v případě, že budou dostupné datové vrstvy o způsobech využití dané oblasti a teplotní data pro celou oblast.

V první kapitole budou obecně popsány geografická data, důležité kartografické pojmy s nimi spojené a také formáty, pomocí kterých je s nimi možné pracovat. V druhé kapitole

bude popsána statistická teorie potřebná k metodám, použitým dále při analýze navržených hypotéz. Ve třetí kapitole budou popsány dostupné datové sady použité v práci, způsob tvorby hypotéz a také shrnuty konkrétní navržené hypotézy ochlazování a oteplování povrchu. V poslední kapitole bude popsán konkrétní postup využití metod ze třetí kapitoly při ověřování navržených hypotéz, bude popsána jejich implementace a dále budou prezentovány výsledky testování hypotéz. Nakonec bude u některých potvrzených hypotéz zkoumán konkrétní vliv jednotlivých prvků, případně jejich poloha v rámci města.



## Kapitola 2

# Geografická data a jejich formát

Tato kapitola se zabývá popisem formátu geografických dat, způsobu jejich uložení a následného zpracování. Jsou zde popsány a porovnány různé formáty a diskutovány jejich výhody a nevýhody. Dále je zde také popsán význam geografického zobrazení a souřadnicového systému.

### 2.1 Geografická data

Geografická data jsou speciálním druhem dat, která popisují polohu objektu na Zemi. Taková data jsou velmi důležitá a užitečná v různých odvětvích jako jsou například navigace, územní plánování, tvorba map nebo analýza prostředí [14]. Proto jsou také hlavním zdrojem pro získání informací o sledované oblasti a pro tvorbu závěrů.

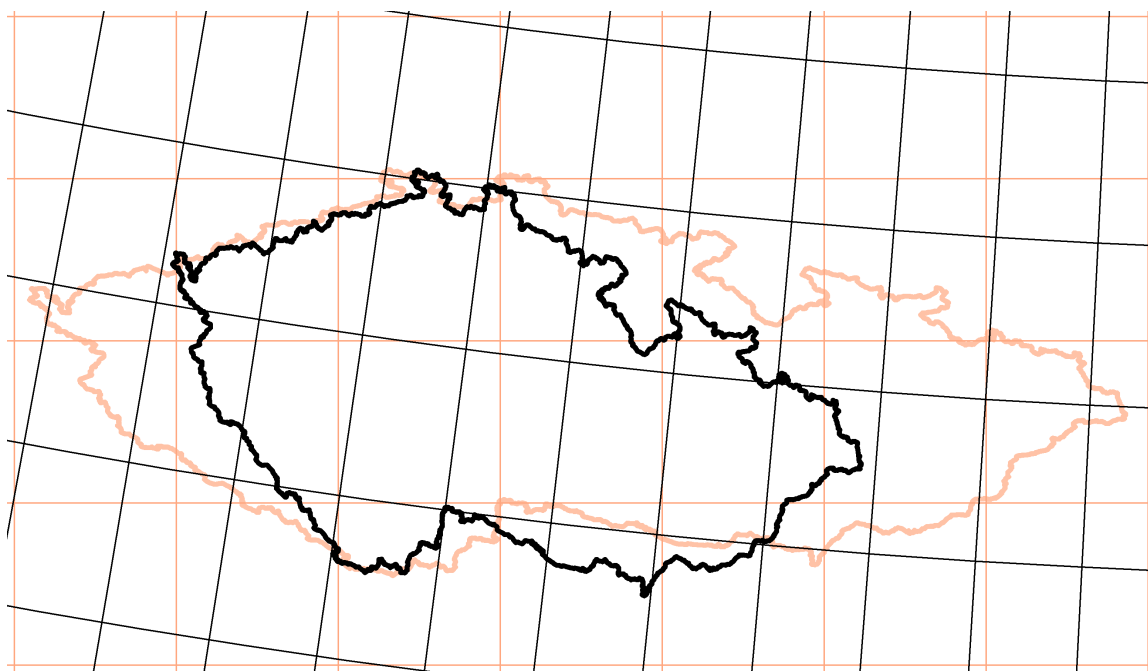
Zdroje a způsoby pořizování dat jsou velmi široké a různorodé. Jedním z nejdůležitějších způsobů při pořizování je využití dálkového průzkumu Země, a to pomocí satelitů, letadel nebo dronů. Další možností jsou měření přímo na zemi. Takto pořízená data lze dále používat jako samostatné datové sady nebo je lze využít k aktualizování či validaci dat pořízených dálkovým průzkumem Země.

Tyto způsoby jsou využívány různými institucemi a společnostmi. Data pořizují například vládní či národní instituce a agentury (Americký národní úřad pro oceán a atmosféru – NOAA, Český úřad zeměměřický a katastrální – ČÚZK, Český statistický úřad – ČSÚ), dále také soukromé komerční společnosti jako např. Google nebo Esri. Existují také projekty, které se snaží pořizovat data cestou open-source jako je OpenStreetMap či projekty založené na crowdsourcingu jako je Waze nebo Mapbox.

### 2.2 Kartografické zobrazení a souřadnicový systém

Důležitou součástí geografických dat a systémů pracujících s nimi je souřadnicový systém a s ním spojené kartografické zobrazení. Zobrazení je způsob, jakým je možné transformovat a zobrazit zemský povrch do rovinné plochy. Zobrazení lze rozdělit podle několika kritérií, a to podle zobrazovací plochy, polohy konstrukční osy a vlastností, které zobrazení zkrusluje nebo zachovává.

Informace v tomto a následujícím odstavci jsou převzaty z [11]. Při použití jedné zobrazovací plochy se jedná o jednoduché zobrazení, konkrétně o azimutální, válcové a kuželové. Pokud nelze konstrukci popsat jednou plochou, hovoří se o nepravém, polykónickém, víceplošném a případně neklasifikovaném zobrazení. Z pohledu polohy osy lze rozdělit zobrazení na osu v polární (shodná se zemskou osou), rovníkové (v rovině rovníku) a horizontální (procházející středem v libovolném směru) poloze. Při zobrazování lze zachovat plochu, v tom případě zobrazení nazývá plochojevné, nebo úhly, kdy se pak zobrazení nazývá úhlojevné, případně lze využít vyrovnávací zobrazení, kde je zkreslení úhlů a ploch v rovnováze. Geografický souřadnicový systém je způsob popisu polohy relativně vzhledem k rovníku a hlavnímu poledníku. Souřadnice jsou poté přepočítávány do souřadnicového systému založeném na zvoleném zobrazení.



Obrázek 2.1: Rozdíl v použití S-JTSK a WGS84 na území ČR (převzato z [https://commons.wikimedia.org/wiki/File:%C4%8Cesk%C3%A1\\_republika,\\_K%C5%99ov%C3%A1k.png](https://commons.wikimedia.org/wiki/File:%C4%8Cesk%C3%A1_republika,_K%C5%99ov%C3%A1k.png))

Pro oblast České republiky, a tudíž i pro tuto práci, jsou zásadní dva geografické souřadnicové systémy, a to konkrétně Systém jednotné trigonometrické sítě katastrální (S-JTSK) a Světový geodetický systém 1984 (World Geodetic System 1984 – WGS-84). S-JTSK je závazným souřadnicovým systémem pro Českou republiku a dle zákona č. 430/2006 Sb.<sup>1</sup> Vychází z Křovákova zobrazení, které je konformním kuželovým zobrazením. Bylo navrženo a je používáno pro svoji přesnost v oblasti geodetických měření v oblasti České republiky. Přesnost S-JTSK se pohybuje od od +14 cm/km až do -10cm/km. Z těchto důvodů je nutné výpočty ploch a vzdáleností provádět na datech, která jsou v tomto souřadnicovém systému. Státní mapová díla jako např. katastrální mapy nebo základní mapy různých měřítek jsou závazně vytvářena v S-JTSK, ale v běžných atlasech a mapách pro prezentaci veřejnosti je využíván WGS-84, a to kvůli tomu, že sever je vychýlen v rozmezí od 4,5° do

<sup>1</sup><https://www.zakonyprolidi.cz/cs/2006-430>

9,5° v závislosti na konkrétní poloze. WGS-84 je na druhou stranu celosvětově používaný souřadnicový systém vytvořený armádou Spojených států z roku 1984. Pracuje s ním globální systém GPS a je to standardní systém pro organizaci NATO. Pro zobrazení je využit zploštělý sféroid [11]. Převody souřadnic mezi těmito systémy jsou velmi složité a není pro tuto práci podstatné se o detailech převodu dále zmiňovat. Převody jsou implementované v knihovnách programovacích jazyků.

## 2.3 GIS software

Pro práci s geografickými daty jsou nezbytné Geografické informační systémy (GIS). Jsou to počítačové programy, umožňující získávání, ukládání, manipulaci analýzu a zobrazení takovýchto dat. Tyto programy lze rozdělit do tří různých kategorií podle způsobu práce s nimi podle prostředí, ve kterém fungují.

Jednou z možností jsou desktopové programy. Mezi nejpoužívanější programy v této oblasti patří ArcGIS Pro od společnosti ESRI. Je to proprietární software a je možné ho využívat pouze na operačním systému Windows<sup>2</sup>. Alternativou k němu je open-source řešení QGIS<sup>3</sup>, které je dostupné i na operačních systémech Unix a macOS. Tyto programy nabízejí širokou škálu nástrojů pro prostorovou analýzu a vizualizaci dat, práce v nich probíhá většinou skrze uživatelské rozhraní a je tedy dostupné i uživatelům, kteří nejsou programátoři. Jsou však náročné na výpočetní a paměťové zdroje a mohou být příliš složité na základní analýzu dat a hledání vztahu dvou vrstev, které bude klíčové pro tuto práci.

Další možností, jak pracovat s geografickými daty, je webový nebo také cloudový GIS software. Mezi zástupce této skupiny programů patří ArcGIS online nebo platforma Google Maps. ArcGIS online nabízí podobné nástroje jako desktopová verze, ale nabízí možnost odlehčené práce oproti desktopové verzi a také jednoduchou možnost sdílení výsledných map a vizualizaci (například i pomocí tzv. příběhových map). Webové prostředí pro práci s geografickými daty nabízí i společnost Google na platformě Google maps. Tato platforma poskytuje několik rozhraní API, jež lze využít pro integrování služeb této platformy do různých aplikací a využití nástrojů například pro geokódování, hledání tras či vyhledávání míst. Tyto služby lze také jednoduše propojit s dalšími nástroji společnosti Google jako Google Sheets nebo Google Cloud Storage. Výhodou těchto řešení je jistě škálovatelnost díky využití cloudového výpočetního výkonu. Omezením může být nabídka pouze základních nástrojů a nemožnost definovat vlastní funkcionalitu pro hledání vztahu mezi datovými vrstvami, dále také dostupnost, případně výkon těchto služeb, které jsou většinou zpoplatněny.

Odlíšným přístupem je využití jazyků běžně používaných pro datovou analýzu (Python, R, Julia) a využití dostupných knihoven pro práci s geografickými daty. V ekosystému jazyka Python existuje mnoho knihoven pro práci s geografickými daty a geometrií obecně. Jednou z nich je GeoPandas, což je rozšíření knihovny Pandas. Pandas umožňuje efektivní práci nad různými datovými sady a GeoPandas přidává možnost pracovat s geometrií dat a jejich polohovou složkou. Geometrické operace jsou možné díky knihovně Shapely, pro načítání vstupních dat je využita knihovna Fiona a pro vykreslování dat jsou využity funkce knihovny matplotlib ve spojení s Contextily pro zobrazení podkladové mapy. Výhodou tohoto přístupu je možnost využívat kromě funkcionality geografických knihoven všechny ostatní funkce jak Pythonu, tak knihovny Pandas. Dále také možnost znovu používat napsané skripty na

---

<sup>2</sup><https://pro.arcgis.com/en/pro-app/latest/get-started/arcgis-pro-system-requirements.htm>

<sup>3</sup><https://www.qgis.org/en/site/>

různé datové sady. Příznivá je jistě i bezplatná dostupnost programovacího jazyka Python a knihoven v něm.

## 2.4 Formáty pro ukládání dat

Pro geografická data potřebujeme speciální formáty dat k reprezentaci jejich polohové části a efektivní práce s nimi. Následuje popis populárních možností těchto formátů.

### 2.4.1 GeoJSON

GeoJSON je otevřený formát pro kódování geografických dat, který je založený na JSON (JavaScript Object Notation). Objekt ve formátu GeoJSON, stejně jako ve formátu JSON, je kolekce dvojic - klíč a hodnota, kde klíče jsou řetězce a hodnoty mohou být řetězce, čísla, pole nebo další objekty [5]. Jednotlivé typy objektů zahrnují:

- Označení pouze geometrie a jejího typu s umístěním v prostoru (Geometry)
- Objekt s dalšími informacemi, který má prostorové souřadnice a typ geometrie (Feature)
- Kolekci objektů z předchozího bodu (FeatureCollection)

Objekt má vždy položku s klíčem `type`, podle kterého se rozlišuje, jestli se jedná přímo o popis pouze geometrie nebo o objekt s dalšími informacemi. Jako hodnotu položky `type` lze uvést buď jeden ze sedmi řetězců označujících typ geometrie nebo `Feature` či `FeatureCollection` [5]. V případě, že se jedná o geometrii, jedinou další položkou jsou souřadnice (`coordinates`) objektu v daném souřadnicovém systému, který v případě GeoJSONU je vždy WGS-84.

```
{
  "type": "Point",
  "coordinates": [102.0, 0.5]
}
```

*Ukázka jednoduché geometrie ve formátu GeoJSON*

Je-li v hodnotě položky `type` uveden řetězec `Feature` nebo `FeatureCollection`, poté může obsahovat položku `geometry`, v jejíž hodnotě je uveden typ geometrie a souřadnice. Volitelně může objekt obsahovat položku `properties`, která je již běžným objektem formátu JSON a vztahují se na ni pravidla pro tento formát. Všechna pojmenování klíčů jsou citlivá na velikost znaků [5].

```

{
  "type": "Feature",
  "geometry": {
    "type": "Polygon",
    "coordinates": [[100.0, 2.5]
  ],
  "properties": {
    "name": "Příklad Polygonu",
    "description": "Toto je popis ukázky polygonu"
  }
}

```

*Ukázka objektu typu **Feature** ve formátu **GeoJSON***

Mezi výhody tohoto formátu patří jeho otevřený formát, čitelnost pro člověka, jelikož jsou veškerá data popsána textově. Dále například jednoduchost a malá režie umožňující jeho použití ve webovém prostředí nebo možnost validovat jeho strukturu. Nevýhodou může být neefektivní ukládání velmi velkých nebo komplexních datových sad, v takovém případě lze využít kompresi pro zmenšení velikosti, jako je např. gzip.

## 2.4.2 Shapefile

Dalším formátem pro uložení dat je proprietární formát společnosti ESRI [9]. Tento formát je starší než výše zmíněný GeoJSON. Shapefile byl vytvořen na začátku devadesátých let, GeoJSON byl oproti tomu formálně definován v roce 2007. Dalším specifikem oproti GeoJSONU je rozdělení jednoho záznamu do více souborů s různými příponami, přičemž všechny tyto soubory musí mít stejný název. Tři z těchto souborů jsou povinné a to konkrétně:

- **.shp** - obsahuje informace o samotné geometrii
- **.shx** - obsahuje index prostorových dat pro zrychlení práce s daty
- **.dbf** - obsahuje dodatečná data o prvcích v datasetu, jako jejich popis, jména, hodnoty

Shapefile nemusí mít a často nemá definovaný souřadnicový systém, ve kterém jsou data uložena. Je možné ho specifikovat v souboru s příponou **.prj**. Pokud záznam tento soubor neobsahuje, je třeba tuto informaci specifikovat, případně odhadnout podle známé lokace a hodnot číselné reprezentace dat.

Shapefile je binární formát, tudíž je pro člověka bez interpretace nečitelný.

## 2.4.3 Formáty založené na XML

Rozšířeným formátem pro výměnu dat je formát XML. Našel své využití i v geografických datech. Jedním z formátů založeným na jazyku XML je formát KML, který je využíván k zobrazování dat v programech jako je Google Earth nebo Google maps [19]. Podobně jako GeoJSON je KML čitelný pro člověka, a tudíž editovatelný v běžném textovém editoru. Formát podporuje běžné typy geografických dat jako body, linie a polygony, ale také například 3D modely nebo obrázky. Data mohou mít také další vlastnosti jako jméno, popis a jiné užitečné informace. Je možné i stylovat barvy, ikonky, popisky prvků. Možnost uchovávat mediální obsah, stylovat v kombinaci s integrací v aplikacích společnosti Google,

která tento formát spravuje a vyvíjí, umožňuje vizuálně atraktivní prezentaci pro uživatele v online prostředí [20].

Dalším formátem založeným na XML je GML, v porovnání s KML se nejedná o jazyk vhodný k vizualizaci dat, prvky uložené v tomto formátu nenesou žádné informace o stylování a vzhledu. Umožňuje však uložit komplexnější geografické informace a jejich topologii než KML [20].

## Kapitola 3

# Metody pro ověřování stanovených hypotéz

Tato kapitola se zabývá obecným popisem statistických metod dále využitých v ověřování hypotéz. Je zde popsán obecný statistický základ a postup pro testování hypotéz, dále konkrétní testy, které nejsou vázány na data s prostorovou složkou, a nakonec testy a metody, které jsou speciálně vytvořené pro geografická data a nacházení vztahů mezi nimi.

### 3.1 Statistické testování hypotéz

Statistické testování je jedním z nejdůležitějších aspektů každého kvantitativního výzkumu. Lze ho využít pro analýzu dat, ověření předpokladů o datech, např. o vztahu jevů na základě dostupných dat, případně vybraného vzorku. Prvním krokem při testování je formulování hypotéz. Hypotéza je tvrzení o náhodné proměnné, které je třeba statisticky ověřit. Rozlišujeme dva typy hypotéz, a to nulovou hypotézu, jejíž platnost ověřujeme a proti ní stojící alternativní hypotézu. Ta může být oboustranná, nebo jednostranná a podle toho lze rozlišit i oboustranné a jednostranné testy. Jednoduchým příkladem může být tvrzení že v určitém faktoru neexistuje signifikantní rozdíl mezi skupinami, což je nulová hypotéza a alternativní hypotéza oproti tomu tvrdí, že existuje signifikantní rozdíl mezi těmito skupinami [12].

Po formulování hypotéz může následovat výběr vzorku z dat. Tento výběr by měl být reprezentativní vzhledem ke všem datům. Vzorkování může být provedeno náhodně nebo například stratifikovaným výběrem [10].

Na těchto datech je spočítána testovací statistika, což je hodnota určující pravděpodobnost platnosti nulové hypotézy. Existuje mnoho druhů testovacích statistik. Jejich výběr záleží konkrétně na tom, jaká data jsou testována nebo jaké otázky jsou kladeny. Lze rozlišit dvě základní skupiny testovacích statistik, a to parametrické a neparametrické [10]. Parametrické testy vyžadují předpoklad normálního rozložení dat nebo by se měly tomuto rozložení alespoň blížit. Příkladem těchto testů je t-test, ANOVA nebo regresní analýza [10]. Na druhou stranu neparametrické testovací statistiky nevyžadují takové předpoklady o datech a jejich rozložení. Mezi neparametrické testy patří Wilcoxonův test pořadí-součet, Kruskal-Wallisův test a chí-kvadrát test. Tyto testy jsou často používány pokud data obsahují odlehlá pozorování, jsou zkreslená nebo velikost vzorku je malá [10].

Další důležitou součástí testování hypotéz je tzv. p-hodnota. Poskytuje míru síly důkazu proti nulové hypotéze. Menší p-hodnota ukazuje na větší sílu důkazů hovořících proti nulové hypotéze. Pokud je p-hodnota menší než zvolená hodnota alfa, typicky 0,05 nebo 0,01, je nulová hypotéza zamítnuta, v opačném případě není nulová hodnota zamítnuta [10].

## 3.2 Statistické testy

T-test je jeden z parametrických testů. Je využíván ke zjištění, jestli se průměry dvou skupin nebo průměr vzorku oproti očekávané hodnotě statisticky významně liší [10]. T-test je založen na tzv. Studentovu t-rozdělení. Toto rozdělení je podobné normálnímu rozdělení, ale je používáno, pokud je vzorek dat malý a jeho rozptyl neznámý [10]. Tvar tohoto rozložení se mění v závislosti na stupních volnosti, jejichž hodnota se určuje podle počtu pozorování. Čím větší stupeň volnosti, tím více se t-rozdělení blíží normálnímu rozdělení [10]. Lze rozlišit dva druhy t-testů, a to jednovýběrový a dvouvýběrový t-test. Jednovýběrový test zkoumá, jestli je vzorek statisticky významně odlišný od očekávaného průměru. Podmínkou je nezávislost mezi vzorky. Párový test je využíván pro dvě skupiny, které spolu souvisí, nepárový pro nezávislé skupiny [10].

Při provádění t-testu je postup totožný s výše popsaným obecným testováním hypotéz. Jsou sesbírána data a poté vypočítány testovací statistiky. Výběr konkrétního testu záleží také na tom, jestli se sobě rovnají rozptyly obou výběrů. Na ověření tohoto faktu je vhodné použít Levenův test. Levenův test spočívá v porovnání mediánů absolutních odchylek jednotlivých pozorování od jejich skupinového mediánu.

$$W = \frac{(N - K)}{(K - 1)} \cdot \frac{\sum_{i=1}^K n_i (Z_{i.} - Z_{..})^2}{\sum_{i=1}^K \sum_{j=1}^{n_i} (Z_{ij} - Z_{i.})^2} \quad (3.1)$$

Výpočet statistiky Levenova testu, kde:

- $N$  je celkový počet pozorování,
- $K$  je počet skupin,
- $n_i$  je počet pozorování ve skupině  $i$ ,
- $Z_{ij}$  je absolutní hodnota odchylky pozorování  $j$  ve skupině  $i$  od mediánu skupiny  $i$ ,
- $Z_{i.}$  je průměrná hodnota  $Z_{ij}$  ve skupině  $i$ ,
- $Z_{..}$  je průměrná hodnota  $Z_{ij}$  přes všechny skupiny.

V případě, že lze na základě Levenova testu říct, že se rozptyly rovnají, je možné použít nepárový t-test:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (3.2)$$

Výpočet statistiky t-testu, kde:



- $\bar{x}_1$   $\bar{x}_2$  značí průměr prvního a druhého vzorku,
- $s_p$  je sdružená standardní odchylka, vypočítána jako:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

kde:

- $n_1$  a  $n_2$  značí velikosti prvního a druhého vzorku
- $s_1$  a  $s_2$  značí standardní odchylky prvního a druhého vzorku

V opačném případě je nutné použít modifikaci nepárového t-testu, známou jako Welchův test nebo Satterthwaitův test:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (3.3)$$

Pokud nelze rozložení dat aproximovat normálním rozložením, je třeba použít tzv. neparametrický test. Neparametrické testy nevyžadují konkrétní předpoklady o rozložení analyzovaných dat. Mezi neparametrické testy porovnávajícími dvě skupiny patří například Wilcoxon Signed-rank využívaný k porovnání párových závislých hodnot nebo Mann Whitney U test, který se používá k porovnání středních hodnot (mediánů) dvou nezávislých skupin [10].

### 3.3 Regrese a prostorová regrese

Regrese je základní statistickou metodou využívanou v rámci analýzy dat při zkoumání vztahu mezi dvěma nebo více proměnnými. Díky této metodě lze odhadnout hodnoty jedné proměnné (závislé) na základě hodnot druhé proměnné (nezávislé), případně skupiny nezávislých proměnných nebo také nalézt vztah mezi těmito proměnnými a sílu tohoto vztahu pomocí koeficientu pro nezávislé proměnné [7].

Existuje mnoho typů regresních modelů, které mají různé využití v závislosti na charakteru dat a cíli analýzy. Mezi nejznámější patří lineární regrese, logistická regrese nebo polynomiální regrese. Lineární regrese je jednoduchý typ regresního modelu, který je ale velmi často používán. Takovýto model popisuje lineární vztah mezi závislou a nezávislou proměnnou nebo skupinou nezávislých proměnných. Lineární regresní model lze vyjádřit jako:

$$Y = X\beta + \epsilon \quad (3.4)$$

kde:

- $Y$  je vektor závislých proměnných
- $X$  je vektor nezávislých proměnných
- $\beta$  je vektor neznámých parametrů, které je nutné odhadnout
- $\epsilon$  je vektor chybových členů

Konkrétní metodou pro lineární regresi je v této práci metoda nejmenších čtverců (ordinary least squares, OLS). Jedná se o široce používanou metodu pro odhad parametrů lineárního regresního modelu. Hlavní princip této metody spočívá v minimalizaci součtu druhých mocnin rozdílů mezi skutečnými a odhadovanými hodnotami závislé proměnné [7]. Přestože je OLS velmi silným nástrojem pro analýzu lineárních vztahů, v oblasti analýzy prostorových dat má své limitace.

Jeden z hlavních problémů je potřeba předpokladu nezávislosti pozorování, což je často u takovýchto dat porušeno kvůli přítomnosti prostorové autokorelace [1]. Tento termín ukazuje na tendenci pozorování, která jsou v prostorové blízkosti, mít podobné hodnoty. Způsobem, jak míru tohoto jevu popsat je Moranovo I (Moran's I) [16]. Při přítomnosti vysoké míry prostorové autokorelace mohou být odhady metody nejmenších čtverců neefektivní a nespolehlivé, stejně tak jako jejich p-hodnoty, což může způsobit mylné závěry [1]. Možným řešením tohoto problému je využití tzv. prostorových regresních modelů. Mezi tyto modely patří model prostorové závislosti (spatial lag model), zohledňující prostorovou autokorelaci v závislé proměnné nebo prostorový chybový model (spatial error model), zohledňující prostorovou autokorelaci v chybových členech nikoliv však samotnou závislou proměnnou. Tyto modely se snaží zohlednit prostorovou autokorelaci a díky tomu produkovat přesnější odhady [1].

Model prostorové závislosti je konkrétním modelem použitým v této práci, který zohledňuje prostorovou autokorelaci. Tento model zohledňuje prostorovou závislost v závislé proměnné prostřednictvím prostorové váhové matice (spatial weights matrix). Matematicky jej lze zapsat jako:

$$Y = \rho WY + X\beta + \epsilon \quad (3.5)$$

kde:

- $Y$  je vektor závislých proměnných
- $\rho$  je koeficient prostorové závislosti
- $W$  je prostorová váhová matice
- $X$  je vektor nezávislých proměnných
- $\beta$  je vektor neznámých parametrů, které je nutné odhadnout
- $\epsilon$  je vektor chybových členů

Model prostorové závislosti zohledňuje prostorovou autokorelaci prostřednictvím prostorové váhové matice (spatial weights matrix), což je způsob, jak popsat prostorové vztahy mezi jednotlivými pozorováními. Častým způsobem, jak konstruovat prostorovou váhovou matici je na základě geografické blízkosti pozorování, mohou však být zahrnuta i jiná kritéria jako např. topologická blízkost v případě dopravních sítí nebo podobnost atributů [2]. Při tvorbě prostorové váhové matice na základě geografické blízkosti mohou být použity různé metody. Nejčastěji používanými jsou tzv. Rook (Věž) a Queen (Královna) kontiguita. Tyto názvy odkazují na způsob, jakým se pohybují ve hře šachy tyto konkrétní figury - věž se pohybuje vodorovně a svisle a královna navíc i diagonálně. Rook kontiguita je založena na přímém sousedství, tudíž za sousední se považují pozorování, která mají společnou hranici. V případě Queen kontiguita se za sousední považují taktéž pozorování mající společnou hranici, ale navíc také pozorování pouze se společným bodem (např. roh) [2]. Díky tomuto tato kontiguita zohledňuje více prostorových vztahů.

Při analýze výstupu regresního modelu je důležité se zaměřit na několik klíčových parametrů, které poskytují výsledky, ale také informace o kvalitě a významnosti modelu. Tyto parametry nám pomáhají posoudit, jak dobře model vysvětluje data a zda jsou zahrnuté proměnné statisticky významné. Výčet nejdůležitějších parametrů [7]:

- $\beta$  Koeficienty regrese – Odhady parametrů regresního modelu, které ukazují vztah mezi nezávislými a závislými proměnnými.
- Standardní chyby koeficientů – Míra variability odhadu koeficientů. Menší standardní chyba znamená přesnější odhad.
- t-statistika – Hodnota t-statistiky pro každý koeficient, která se používá k testování hypotézy, že koeficient je roven nule.
- p-hodnoty – Určují maximální hodnotu hladiny významnosti pro kterou je hodnota dané nezávislé proměnné statisticky významná.
- Koeficient determinace ( $R^2$ ) – Hodnota vyjadřující jaký podíl celkové variability závislé proměnné je vysvětlen nezávislými proměnnými v modelu. Čím je  $R^2$  bližší 1, tím lépe model vysvětluje data.
- Přizpůsobený (adjusted) koeficient determinace – Upravená verze  $R^2$ , která zohledňuje počet nezávislých proměnných ve srovnání s počtem pozorování.

## Kapitola 4

# Návrhy hypotéz a způsobu jejich ověření

V této kapitole jsou popsány vyslovené hypotézy pro ochlazování či oteplování oblastí na území města Brna a jakým způsobem jsou definovány pro jednotlivé druhy dat. Jsou zde popsány dostupné datové sady, jejich formát a zpracování a dále jejich využití při definování hypotéz.

### 4.1 Způsob tvorby hypotéz

Hlavním cílem hypotéz je odpovědět na otázku co, jakým způsobem a jak silně může ovlivňovat teplotu na území města Brna. Hypotézy budou tvořeny vzhledem k dostupným datovým sadám a k potenciální možnosti daného jevu ovlivňovat teplotu ve svém okolí. Hlavní způsob pro formulování hypotéz je otázka, jestli jsou výskyty prvků datové sady, případně jejich hodnoty ve vztahu s oblastmi vyšší, případně nižší teploty. V souvislosti s takto formulovanou otázkou se lze zabývat tím, jak silný tento vztah je a jestli je statisticky významný s využitím metod popsaných v předchozí kapitole. Data jsou ve většině případů polygonového charakteru, intenzita dopravy je liniová vrstva a zelené střechy jsou bodová vrstva. Při zpracování těchto vrstev je potřeba je převést na polygony pomocí vytvoření obálky okolo nich.

### 4.2 Dostupné datové sady

Datové sady, které budou v práci využity, jsou ve všech případech volně dostupné na portálu Magistrátu města Brna na webovém portálu [data.brno.cz](http://data.brno.cz). Je možné je získat v různých běžných formátech popsaných v kapitole 2, jako jsou GeoJSON, KML, Shapefile a dalších. Jako formát pro stažení dat byl zvolen GeoJSON, kvůli možnosti stáhnout datovou sadu jako jeden soubor narozdíl od Shapefile. Navíc pro výpočty pomocí statistických metod není důležitá prezentace ve webovém prostředí, kterou nabízí např. KML. Přehled využitých datových sad je popsán v následující tabulce, podrobnější popis každé datové sady následuje níže.

### 4.2.1 Teplota povrchu

Nejdůležitější datovou sadou je teplota povrchu<sup>1</sup>. Vizualizována je pomocí teplotní mapy na obrázku 4.1. Tato data byla pořízena 9. 9. 2020 v 9:44. Průměrná teplota vzduchu na stanici Brno-Tuřany v tento den byla 18,4 °C, maximální pak 23,7 °C, tyto údaje byly získány z dat ČHMÚ<sup>2</sup>. Snímek byl pořízen družicí Landsat 8. Tato družice nese dva moduly pro pořizování snímků povrchu Země. Pro teplotu povrchu byl využit Teplotní infračervený senzor – Thermal Infrared Sensor (TIRS), který umožňuje vytvářet snímky ve dvou rozlišeních a to 15 a 30 metrů podél 185kilometrového záběru s využitím kvantové fyziky<sup>3</sup>. Dle popisu byl původní rasterový formát dat vektorizován do podoby spojitých nepravidelných polygonů reprezentujících teplotu povrchu. Polygon vždy reprezentuje jednu hodnotu teploty, teploty jsou zde uvedené jako desetinná čísla s přesností na jedno desetinné místo. Průměrná teplota povrchu v celé oblasti v danou chvíli byla asi 22,3 °C. Jelikož byl snímek pořízen v dopoledních hodinách a mimo hlavní letní měsíce, lze předpokládat, že by teplota v pozdějších hodinách po vyhřátí povrchu slunečním zářením nebo v horkých letních dnech byla výrazně vyšší. Tato datová sada bude sloužit jako základ pro zkoumání faktorů, které mohou mít na teplotu povrchu vliv. Je to jediná dostupná celistvá teplotní sada – mapa pro město Brno. Snímek se dle dokumentace aktualizuje jednou ročně, ale novější data než právě z roku 2020 dostupná nejsou. Dle komunikace s brněnským magistrátem, bohužel nejsou dostupná žádná data ze snímků z předchozích let. Na portálu data.brno.cz je dostupná také bodová vrstva teplot z čidel Tepláren Brno, a.s.<sup>4</sup>, která je aktualizována každých 15 minut. Takováto bodová data nejsou ale vhodná pro prostorovou analýzu, jelikož nepřináší detailní informaci o teplotě pro celou oblast. Bylo by možné data extrapolovat, pro vytvoření celistvé vrstvy, ale užitečnost takovéto operace vzhledem k velké prostorové variabilitě teploty by nebyla dobrá, navíc ostatní vrstvy také nemají historické verze dat.

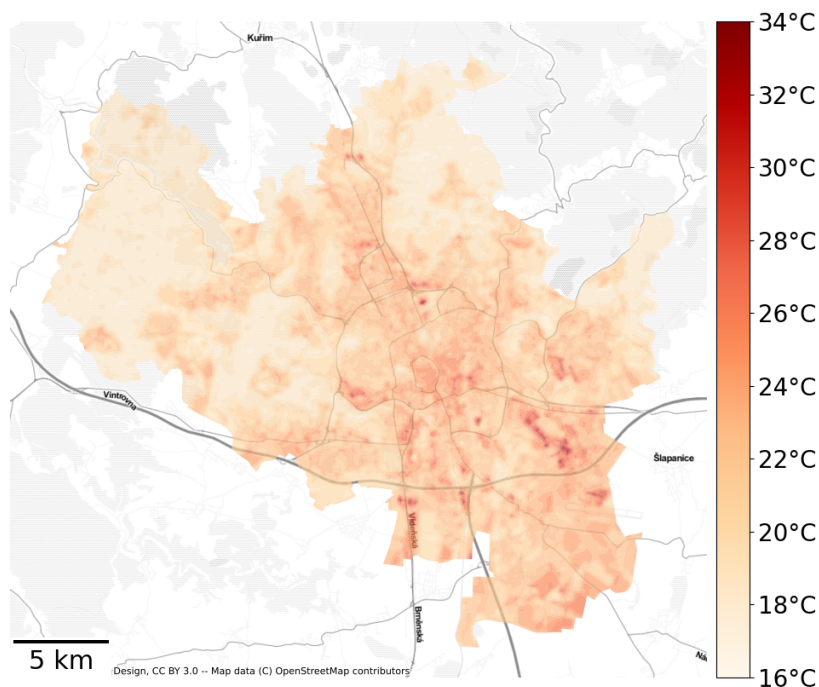
---

<sup>1</sup><https://data.brno.cz/datasets/mestobrnno::teplota-povrchu-surface-temperature/about>

<sup>2</sup><https://www.chmi.cz/historicka-data/pocasi/denni-data/data-ze-stanic-site-RBCN#>

<sup>3</sup><https://www.usgs.gov/landsat-missions/landsat-8>

<sup>4</sup><https://data.brno.cz/datasets/mestobrnno::pr%C5%AFm%C4%9Brn%C3%A9-aktu%C3%A1ln%C3%AD-teploty-vzduchu-average-air-temperatures/about>



Obrázek 4.1: Vykreslené hodnoty teplot povrchu z datové sady 1

#### 4.2.2 Industriální areály, brownfieldy a skládky

Industriální areály<sup>5</sup> a skládky<sup>6</sup> pocházejí z aplikace Georizika, kterou vytvořil odbor životního prostředí Magistrátu města Brna. Obsahuje mnoho různých datových sad, které dávají možnost udělat si představu o georizicích na území města. V této práci budou využity tři potenciálně zajímavé vrstvy, které by mohly mít vliv na teplotu povrchu, a to konkrétně vrstva industriálních areálů, skládek a antropogenních uloženin. Brownfieldy<sup>7</sup> jsou samostatně vytvořená datová sada oddělením rozvojových lokalit magistrátu. Tyto vrstvy jsou vizualizovány společně s plochou, kterou ve městě zabírají na obrázku 4.2

Industriální areály jsou polygonová vrstva, ve které jsou zaznamenané areály, ve kterých bylo nebo je nakládáno se škodlivými látkami pro životní prostředí. V datech je rovněž uveden název, případně majitel areálu, adresa, stručný popis využití areálu a další údaje. Při ověřování vlivu těchto areálů může být dále vhodné kromě vlivu areálů obecně studovat i míru vlivu jednotlivých areálů. Teplo produkované průmyslovou výrobou by také mohlo zvyšovat teplotu areálu a jeho okolí [23].

**Hypotéza 1** *Industriální areály přispívají ke zvyšování teploty na svém území*

Skládky mohou zvyšovat teplotu ve svém blízkém okolí, a tak přispívat ke zvýšení teploty v širším okolí, a to díky vypouštění skleníkových plynů a emisí zachycujících teplo. Při rozpadání odpadu, především organického se může uvolňovat metan. Ten může být až několikrát účinnější při zachycování tepla než oxid uhličitý [3]. Skládky mohou často také

<sup>5</sup><https://data.brno.cz/datasets/mestobrna::georizika-are%C3%A1ly-geological-risks-industrial-areas/about>

<sup>6</sup><https://data.brno.cz/datasets/mestobrna::georizika-sk%C3%A1dky-a-antropogenn%C3%AD-ulo%C5%BEeniny-geological-risks-landfills-and-anthropogenic-deposits/about>

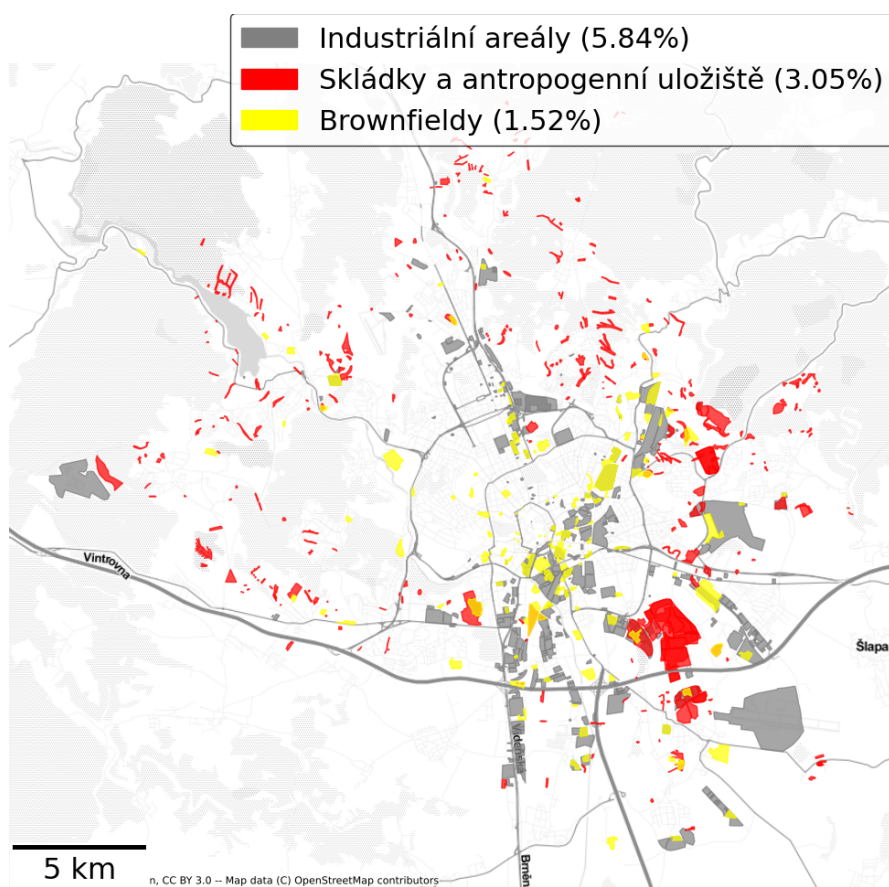
<sup>7</sup><https://data.brno.cz/datasets/mestobrna::brownfields-rozvojov%C3%A9-lokality/about>

obsahovat materiály tmavšího zbarvení, tudíž lépe absorbující sluneční světlo a přispívající ke zvýšení efektu tepelného ostrova.

### **Hypotéza 2** *Skládky přispívají ke zvyšování teploty na svém území*

Podobně brownfieldy, dříve průmyslové objekty, které jsou nyní opuštěny nebo nevyužívány, mohou ovlivňovat teplotu na svém území. Mnoho z těchto areálů je charakterizováno materiály jako beton, asfalt nebo ztuhlá půda, které mají vysokou tepelnou kapacitu a absorbují sluneční světlo, což také může přispívat k efektu městského tepelného ostrova [17]. Navíc brownfieldy často postrádají vegetaci, což znamená, že je zde méně potenciálního stínění, které by pomohlo ochladit oblast.

### **Hypotéza 3** *Brownfieldy přispívají ke zvyšování teploty na svém území*



Obrázek 4.2: Mapa industriálních areálů 5, skládek 6 a brownfieldů 7 na území města Brna s údajem kolik procent plochy města reprezentuje

### **4.2.3** **Využití půdy**

Pro informace o využití půdy jsou kromě dat na portálu data.brno.cz<sup>8</sup> dostupná také data z RÚIAN (Registr územní identifikace, adres a nemovitostí) ve spojení s OpenStreetMaps<sup>9</sup>.

<sup>8</sup><https://data.brno.cz/datasets/mestobrna:vyu%C5%BEit%C3%AD-p%C5%AFdy-land-cover/about>

<sup>9</sup><https://wiki.openstreetmap.org/wiki/Cs:R%C3%9AIAN#Parcely>

Při rozhodování, který datový zdroj využít, bylo přihlédnuto k datu vytvoření datové sady na portálu data.brno.cz, která je časově blízko datu pořízení teplotních dat (červen 2020 - září 2020). Vzhledem k účelovému pořízení této datové sady přímo pro město Brno pomocí družice lze potenciálně očekávat větší přesnost dat. Využity tedy budou data z portálu data.brno.cz

Data o využití půdy jsou strukturována jiným způsobem než předchozí datové sady. Nerepresentují jednotlivé prvky určitého jevu, ale pokrývají celou oblast města, rozložení jednotlivých kategorií lze vidět na obrázku 4.3. Město je rozděleno do kategorií, a to konkrétně: zastavěné území, doly a lomy, travní porost, les, orná půda, křoviny a vodní plocha.

**Zastavěná území** mohou zvyšovat teploty prostřednictvím efektu městského tepelného ostrova. Stavební materiály, jako je beton a asfalt, absorbují a udržují teplo a zároveň nedostatek zeleně snižuje schopnost prostředí přirozeně se ochlazovat [17].

**Těžba v lomech** může způsobit lokální změny teploty v důsledku odstraňování vegetace a odhalování půdy či hornin, které mají jinou tepelnou kapacitu než původní vegetace. Mění se krajina může také ovlivnit proudění větru a odpařování.

**Hypotéza 4** *Využití půdy jako zastavěné území nebo lomy ve svých oblastech přispívá ke zvýšení teploty*

**Travnaté porosty**, jako jsou louky a pastviny, mohou mírně snižovat teploty prostřednictvím evapotranspirace, kdy rostliny vypouštějí vodu zpět do atmosféry. Travníky také odrážejí více slunečního záření než nezelené povrchy, jako je beton nebo asfalt [18].

**Lesy** mají tendenci snižovat teploty prostřednictvím evapotranspirace a stínění, což zabraňuje přímému ohřevu zemského povrchu slunečním zářením. Kromě tohoto efektu lesy pomáhají zmírňovat efekt globálního oteplování pohlcováním a ukládáním oxidu uhličitého [4].

**Křoviny** mohou mít mírně chladicí účinek, protože poskytují stín a zprostředkovávají evapotranspiraci, i když míra tohoto účinku je určitě menší než u lesů nebo travnatých porostů. Přesto mohou křoviny pomoci snížit teploty ve srovnání s nezelenými povrchy, jako je beton nebo asfalt [18].

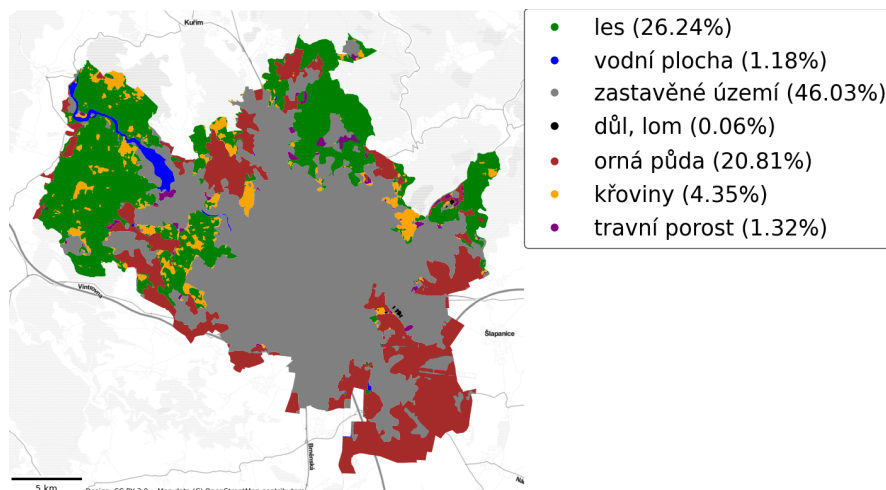
**Hypotéza 5** *Využití půdy jako zelená plocha (les, křoviny, travní porost) přispívá ke snižování teploty na svém území*

**Orná půda** může mít smíšený vliv na teploty, v závislosti na typu plodiny a zemědělských postupů. Evapotranspirace plodin může pomoci snížit teploty, ale odlesňování a ztráta přirozené vegetace kvůli zemědělskému využití mohou vést ke zvýšení teplot [15].

**Vodní plocha** Vodní plochy, jako jsou vodní toky (Svratka a Ponávka) a nádrže nebo jezera (Brněnská přehrada), mohou mít chladicí účinek na okolní prostředí prostřednictvím odpařování a zvýšené vlhkosti vzduchu. Vodní plochy také absorbují a udržují teplo jinak než pevninské povrchy, což může mít za následek lokální změny teploty a mikroklimatu [8].

**Hypotéza 6** *Využití půdy jako vodní plocha přispívá ke snižování teploty na svém území*





Obrázek 4.3: Mapa druhů využití půdy z datasetu 8 s údajem kolik procent plochy města reprezentuje

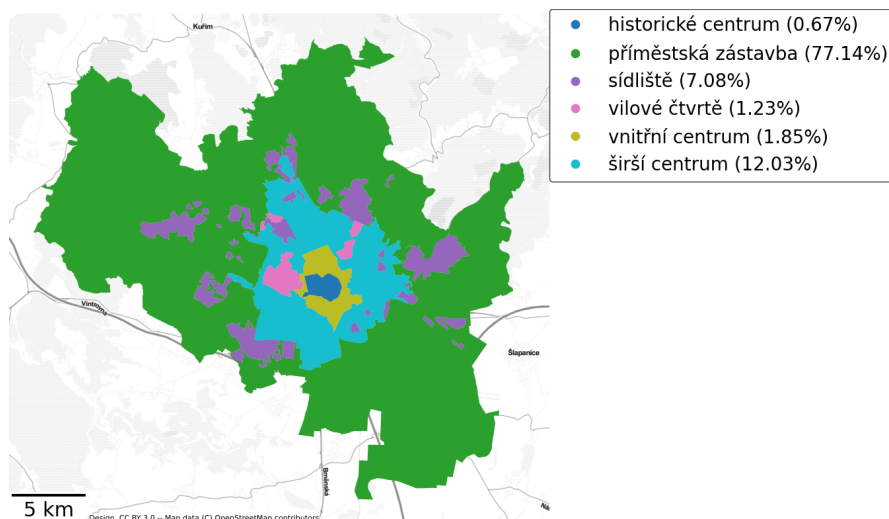
#### 4.2.4 Druh městské zástavby

Podobně jako u předchozí datové sady využití půdy jsou dostupná data z portálu data.brno<sup>10</sup> a z RÚIAN<sup>11</sup>, kde je zaznamenáno jaký druh stavby se nachází na jaké parcele. Při rozhodování zde bylo přihlédnuto k účelovému vytvoření sady magistrátem, tudíž potenciálně lepší kvalité dat než obecný zdroj pro celou Českou republiku. Data z RÚIAN jsou také velmi podrobná, obsahují velké množství druhů staveb. Data z portálu data.brno.cz obsahují 6 kategorií druhů zástavby, se kterými se bude lépe pracovat než s jednotlivými parcelami, které by oblast rozdělily na velké množství malých oblastí. Tato vrstva tedy rozděljuje město na 6 různých druhů zástavby, jejichž rozložení je vizualizováno na obrázku 4.4. Více zastavěné části by mohly přispívat k vyšším teplotám díky většímu zastoupení materiálů jako beton, asfalt apod. [17]. Sídliště, vnitřní a historické centrum by mohly díky tomu mít na svém území vyšší teploty povrchu. Příměstská zástavba hlavně díky tomu, že obsahuje okrajové, méně obydlené části města, by mohla přispívat k ochlazování. U ostatních částí nelze předem odhadnout, jakým způsobem teplotu ovlivní.

**Hypotéza 7** *Sídliště, vnitřní a historické centrum coby druh městské zástavby přispívá ke zvyšování teploty na svém území*

<sup>10</sup><https://data.brno.cz/datasets/mestobrno:m%C4%9Bstsk%C3%A1-z%C3%A1stavba-charakter-city-built-up-area-types/about>

<sup>11</sup><https://wiki.openstreetmap.org/wiki/Cs:R%C3%A1IAN#Parcely>



Obrázek 4.4: Mapa charakteru městské zástavby z datasetu 10 s údajem kolik procent plochy města reprezentuje

#### 4.2.5 Významné plochy zeleně

Datová sada obsahuje informace o nejvýznamnějších plochách zeleně ve městě Brně<sup>12</sup>, které jsou součástí „Seznamu ploch nejvýznamnější zeleně v městě Brně, jejichž plošný rozsah nesmí být zmenšen“, a tvoří přílohu obecně závazné vyhlášky č. 15/2007 o ochraně zeleně ve městě Brně. Tyto plochy zahrnují nejvýznamnější městské parky, menší parky či zelená veřejná prostranství a větší plochy městské zeleně, jako jsou izolační nebo sídlištní zeleně, které hrají klíčovou roli v celkovém území města. Nejvýznamnější městské parky a zelená veřejná prostranství poskytují stín, zprostředkovávají evapotranspiraci a tvoří zelené koridory, které mohou pomoci snižovat teploty ve městě [6]. Menší parky a veřejná prostranství mohou mít podobný chladicí účinek, i když v menší míře. Přítomnost zeleně na sídlištních může taky pomáhat se snižováním teplot na nich.

**Hypotéza 8** *Významné plochy zeleně přispívají ke snižování teploty na svém území*

#### 4.2.6 Intenzita dopravy

Jedná se o liniovou datovou vrstvu zpracovávanou společností Brněnské komunikace a.s.<sup>13</sup>. Ukazuje každoroční intenzitu dopravy na vybraných úsecích silnic od roku 2016 do roku 2021. Intenzita je vyjádřena jako počet motorových vozidel v tisících za 24 hodin. Hodnoty jsou udávány jak pro osobní automobily, tak pro nákladní vozidla. Větší provoz na silnicích by mohl mít vliv na zvýšení teploty silnice a jejího okolí, ať už kvůli teplu, které auta vytvářejí motorem a spalováním paliva nebo kvůli pohybu vozidel po silnici. Jako ulice s vyšší intenzitou dopravy byly použity ulice, kde byla intenzita dopravy vyšší než 0,75 kvantil.

**Hypotéza 9** *Ulice s vyšší intenzitou dopravy budou zvyšovat teplotu ve svém okolí*

<sup>12</sup><https://data.brno.cz/datasets/mestobrna::v%C3%BDznamn%C3%A9-plochy-zelen%C4%9B-major-parks-and-green-areas/about>

<sup>13</sup><https://data.brno.cz/datasets/mestobrna::intenzita-dopravy-intenzita-vozidel-vehicle-traffic-intensity/about>

#### 4.2.7 Zelené střechy

Zelené střechy jsou bodovou vrstvou<sup>14</sup>. Pouze 119 ze střech bylo realizováno před rokem 2020, ze kterého jsou data o teplotách povrchu, tudíž zbytek datové sady nelze použít. Vegetace na zelené střeše může fungovat podobně jako zeleň v parcích, tedy zvyšovat evapotranspiraci a snižovat absorpci tepla. Oproti jiným materiálům využívaným na střechy má také zeleň menší odrazivost slunečního záření [22]. Všechny tyto faktory mohou přispívat ke snížení okolní teploty. Problémem u této vrstvy je neznalost velikosti střechy, jelikož vrstva je bodová.

**Hypotéza 10** *Zelené střechy přispívají ke snižování teploty na své ploše*

---

<sup>14</sup><https://data.brno.cz/datasets/mestobrnno::realizovan%C3%A9-zelen%C3%A9-st%C5%99echy-green-rooftops-/about>

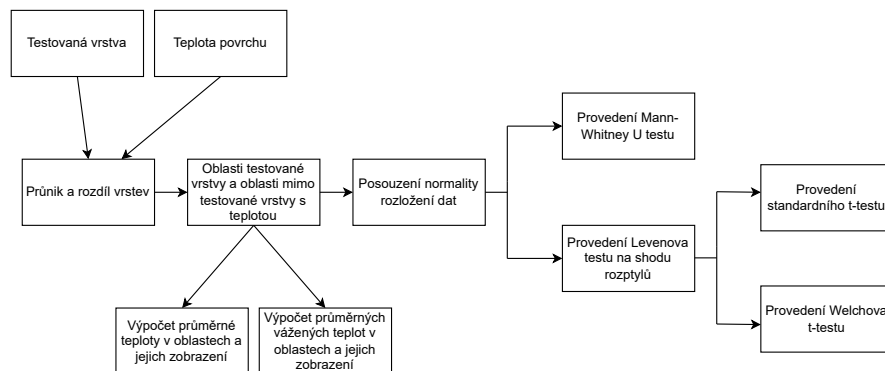
## Kapitola 5

# Implementace ověřování hypotéz a jejich testování

V této kapitole bude popsáno, jakými konkrétními způsoby bude využito metod (statistické testy, regresní analýza) popsaných v kapitole 3 a způsob transformování dat do potřebné podoby. Poté budou prezentovány pomocí grafů a tabulek výsledky obou metod u jednotlivých vrstev.

### 5.1 Postup při využití statistických testů a porovnání průměrů

Při využití statistických testů bude cílem testovat nulovou hypotézu, že průměr teplot v oblasti jevu dané datové sady je statisticky signifikantně vyšší, případně nižší než průměr teplot mimo tyto oblasti (např. testování zda průměr teplot v oblastech industriálních areálů je vyšší než mimo ně). Všechny statistické testy budou prováděny se standardní hodnotou významnosti 0,05. Postup při testování je znázorněn na obrázku 5.1 a dále také popsán slovně.



Obrázek 5.1: Diagram postupu při testování hypotéz pomocí statistických testů

Prvním krokem při využití této metody je vytvoření pomocného GeoDataframe, který reprezentuje průnik testované vrstvy dat a vrstvy teplot. K tomu je vytvořen GeoDataframe reprezentující oblasti mimo oblast testovaného jevu, vytvořený pomocí prostorové operace rozdíl. Dále je spočítána vážená průměrná teplota v obou těchto oblastech, přičemž jako

váha je použita plocha jednotlivých polygonů. Před spočítáním konkrétní testovací statistiky pro testování hypotézy je nutné zjistit, zda rozložení teplot obou oblastí splňují podmínku normality. Pokud nemají data normální rozložení, je třeba použít neparametrický test, Mann-whitney U. V opačném případě je třeba prozkoumat, zda mají obě skupiny stejný rozptyl. Lze využít Levenův test 3.1, případně F-test. Výsledek těchto testů je důležitý pro rozhodnutí, jestli je potřeba využít standardní t-test 3.2, vyžadující stejné rozptyly u obou skupin, nebo tzv. Welchův t-test 3.3, který tuto podmínku nevyžaduje. Dle výsledků lze rozhodnout, že zamítáme nulovou hypotézu, tudíž že teploty v oblastech zkoumaného jevu nejsou dostatečně statisticky vyšší než v oblastech, kde se jev nevyskytuje. To všechno na zvolené hladině významnosti 0,05.

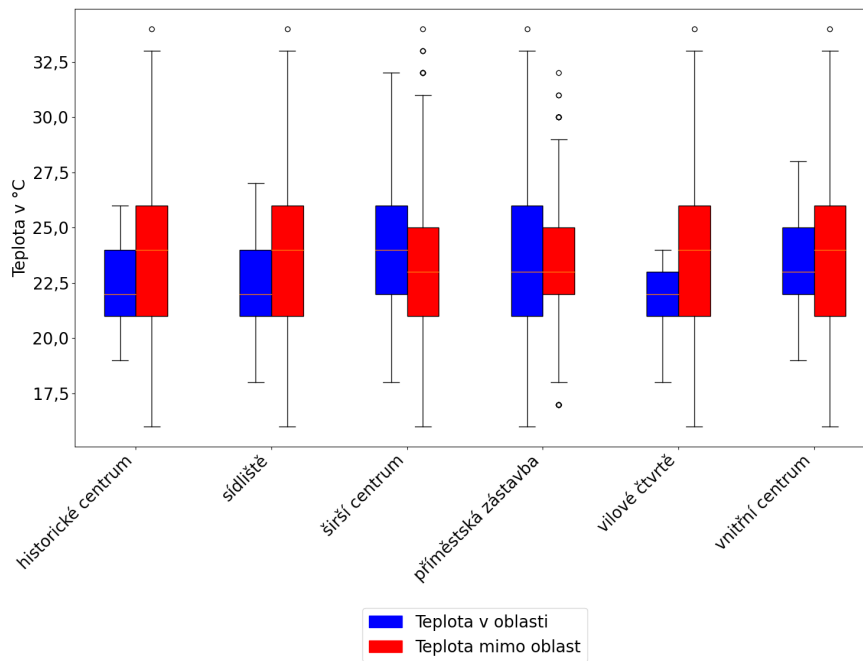
## 5.2 Výsledky testování hypotéz pomocí statistických testů

V následující podkapitole jsou prezentovány výsledky testování jednotlivých datových sad. Jsou zde uvedené porovnání vážených průměrů teplot v oblasti, přičemž jako váha je použita plocha dané oblasti, a dále jsou zde uvedené p-hodnoty provedených testů a jejich textové shrnutí.

### 5.2.1 Městská zástavba

Tabulka 5.1: Porovnání vážených průměrů u jednotlivých druhů městské zástavby

Kategorie	Vážený průměr v oblasti ve °C	Vážený průměr mimo oblast ve °C	Rozdíl
historické centrum	23,24	20,98	2,26
sídlíště	22,26	20,90	1,36
širší centrum	22,70	20,76	1,94
příměstská zástavba	20,53	22,57	-2,04
vilové čtvrtě	21,73	20,99	0,74
vnitřní centrum	23,28	20,95	2,33



Obrázek 5.2: Porovnání nevážených průměrů teplot v různých oblastech městské zástavby

Z datové sady druhů městské zástavby byly otestovány všechny dostupné druhy, které pokrývají celou oblast měst. Jednotlivé druhy v sobě zahrnují mnoho různých dalších faktorů, které mohou teplotu ovlivňovat. Z prvotního porovnání vážených průměrů v oblasti dané zástavby a mimo něj v tabulce 5.1 lze pozorovat, že kromě oblasti příměstské zástavby je ve všech ostatních oblastech teplota vyšší než mimo ně, což je způsobené hlavně velikostí této oblasti. Avšak pokud by nebyla brána v potaz plocha oblastí jako váha pro průměr jako na obrázku 5.2, což v samotném statistickém testování brána není, tak lze pozorovat spíše opačný trend. Signifikantnost těchto rozdílů byla statisticky otestována a výsledky jsou v tabulce 5.2 níže.

Tabulka 5.2: Výsledky statistických testů u jednotlivých druhů městské zástavby

Vrstva	Využitý test	P-hodnota
historické centrum	Mann-Whitney U	$2,563 \times 10^{-02}$
vilové čtvrtě	Mann-Whitney U	$6,184 \times 10^{-05}$
vnitřní centrum	Mann-Whitney U	$3,506 \times 10^{-01}$
sídliště	Welchův t-test	$1,120 \times 10^{-24}$
širší centrum	Welchův t-test	$1,386 \times 10^{-16}$
příměstská zástavba	Welchův t-test	$2,269 \times 10^{-01}$

U třech druhů zástavby, jejichž rozložení nebylo blízké normálnímu, byl proveden jednostranný Mann-Whitney U test, který testoval nulovou hypotézu, že mediány těchto dvou oblastí jsou stejné a alternativní hypotézu, že medián mimo oblasti zástavby je vyšší. U historického centra a vilových čtvrtí nulovou hypotézu zamítáme a teplota mimo tyto oblasti je signifikantně vyšší, u vnitřního centra nulovou hypotézu nezamítáme.

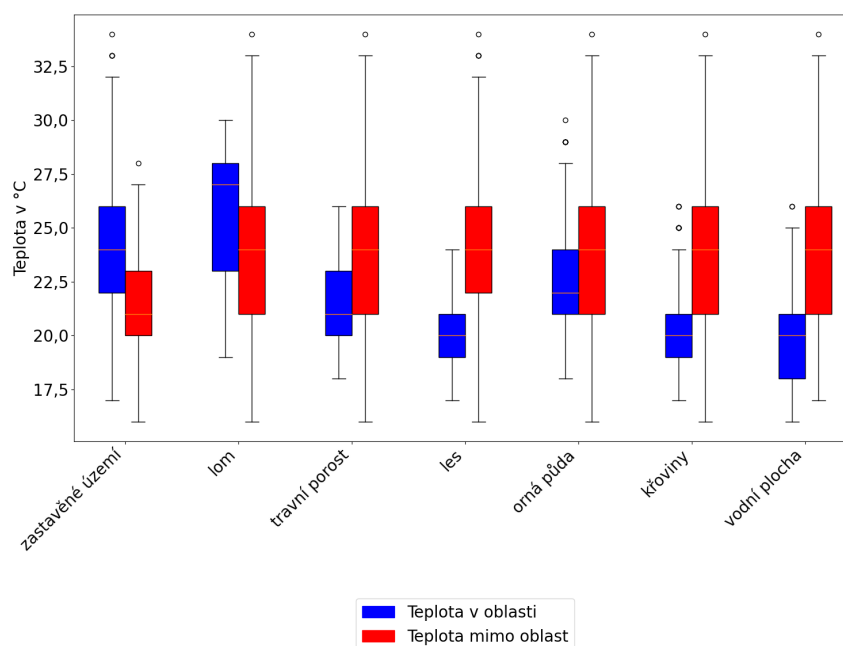
Dle výsledků Levenova testu rozptylů byl na zbylé oblasti využit jednostranný Welchův t-test, který testuje stejnou nulovou hypotézu, ale na rozdíl od Mann-Whitney U testuje průměry místo mediánu. U sídlišť a širšího centra zamítáme nulovou hypotézu, tudíž teplota v oblasti je signifikantně vyšší a u příměstské zástavby nulovou hypotézu nezamítáme a teplota není signifikantně vyšší mimo oblast.



## 5.2.2 Využití půdy

Tabulka 5.3: Porovnání vážených průměrů u jednotlivých využití půdy

Kategorie	Vážený průměr v oblasti ve °C	Vážený průměr mimo oblast ve °C	Rozdíl
zastavěné území	22,13	19,88	2,25
lom	25,55	20,99	4,56
travní porost	20,70	21,00	-0,30
les	18,49	21,98	-3,49
orná půda	22,03	20,70	1,34
křoviny	19,34	21,08	-1,74
vodní plocha	18,05	21,03	-2,98



Obrázek 5.3: Porovnání nevážených průměrů teplot v různých oblastech využití půdy

Z datové sady způsobu využití půdy byly stejně jako u městské zástavby otestovány všechny způsoby. Vážené průměry jsou prezentovány v tabulce 5.3 a rozložení nevážených průměrů je vizualizováno na obrázku 5.3. Velmi výrazné rozdíly lze na první pohled pozorovat u dolů a lomů a také u lesů a vodních ploch. U jediné orné půdy lze pozorovat změnu v tom, která z teplot je vyšší v případě že není započítána plocha oblastí jako váha. Použité testy a jejich p-hodnoty jsou opět v tabulce níže 5.4.

Tabulka 5.4: Výsledky statistických testů u jednotlivých využití půdy

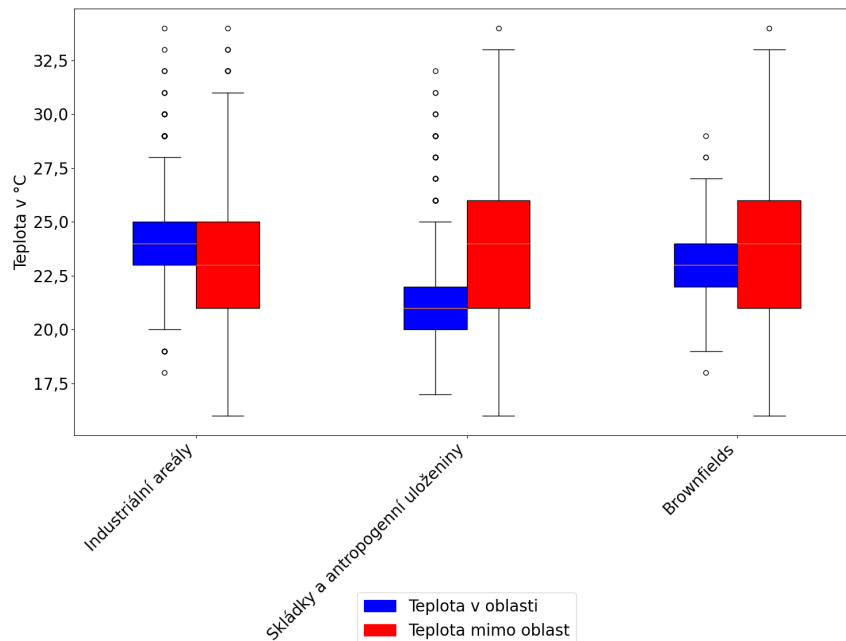
Vrstva	Využitý test	P-hodnota
vodní plocha	Mann-Whitney U	$1,237 \times 10^{-24}$
křoviny	Mann-Whitney U	$1,047 \times 10^{-71}$
orná půda	Mann-Whitney U	$1,875 \times 10^{-24}$
lom	Mann-Whitney U	$8,412 \times 10^{-07}$
zastavěné území	Welchův t-test	$7,264 \times 10^{-185}$
travní porost	Welchův t-test	$3,047 \times 10^{-38}$
les	Welchův t-test	$6,842 \times 10^{-285}$

Část způsobů využití půdy dobře aproximuje normální rozložení, část ne (např. u vodních ploch nebo dolů a lomů je to způsobeno malým počtem těchto jevů). U všech testů je p-hodnota nižší než 0,05, zamítáme tedy nulové hypotézy. Tudíž v zastavěném území a dolech nebo lomech je signifikantně vyšší teplota než mimo ně a ve všech ostatních oblastech využití půdy je teplota signifikantně nižší.

### 5.2.3 Industriální areály, skládky a brownfieldy

Tabulka 5.5: Porovnání vážených průměrů u industriálních areálů, skládek a brownfieldů

Kategorie	Vážený průměr v oblasti ve °C	Vážený průměr mimo oblast ve °C	Rozdíl
Industriální areály	23,78	20,83	2,95
Skládky a antropogenní uložení	20,66	20,99	-0,33
Brownfieldy	22,41	20,97	1,44



Obrázek 5.4: Porovnání vážených průměrů u industriálních areálů, skládek a brownfieldů

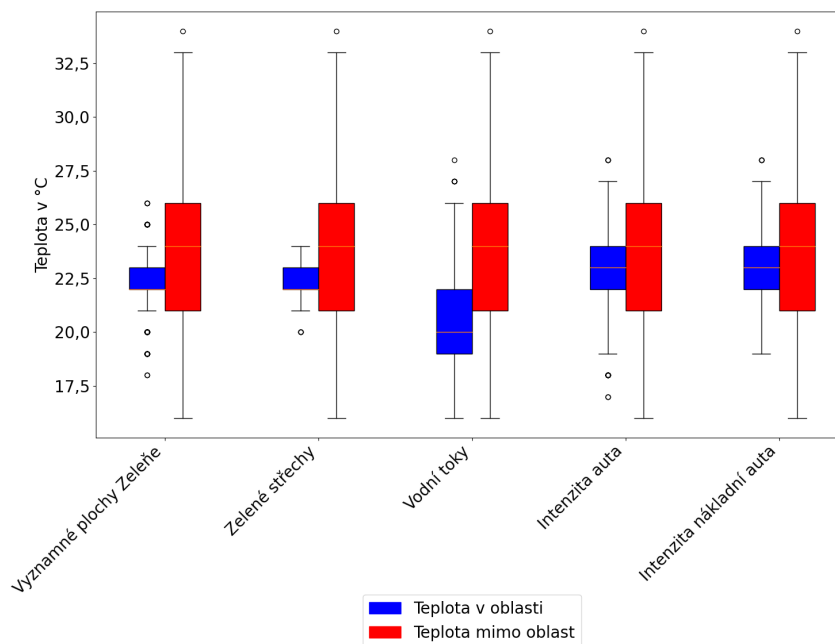
Oproti předchozím sadám, které pokrývají celé území měst a definují všude jednu z kategorií jsou sady z programu Georizika menšími oblastmi, které budou sdílet velkou část mimo svůj prostor. Vážený průměr zobrazený v tabulce 5.5, má vliv na to který průměr je větší pouze u brownfieldů, u zbylých dvou je vliv pouze ve velikosti rozdílu průměru teplot. Ve statistických testech je opět použit medián nebo nevážený průměr 5.4. P-hodnoty a použité testy pro jednotlivé vrstvy jsou prezentovány v tabulce 5.6 níže.

Tabulka 5.6: Výsledky statistických testů u industriálních areálů, skládek a brownfieldů

Vrstva	Využitý test	P-hodnota
Skládky a antropogenní uložení	Mann-Whitney U	$6,223 \times 10^{-113}$
Industriální areály	Welchův t-test	$7,081 \times 10^{-38}$
Brownfields	Welchův t-test	$1,410 \times 10^{-13}$

Rozložení teplot industriálních areálů a brownfieldů dobře aproximuje normální rozložení, ale nelze hovořit o rovnosti rozptylů, proto je použitý Welchův t-test a u vrstvy skládek je použit Mann-Whitney U test. P-hodnoty jsou výrazně nižší než běžná hranice 0,05. Zamítáme tedy nulové hypotézy a rozdíly viditelné na grafu nevážených průměrů jsou statisticky signifikantní.

## 5.2.4 Různé další datové sady



Obrázek 5.5: Porovnání průměrů teplot v různých dalších datových sadách

Tabulka 5.7: Výsledky statistických testů u dalších datových sad

Vrstva	Využitý test	P-hodnota
Vodní toky	Mann-Whitney U	$3,762 \times 10^{-143}$
Zelené střechy	Mann-Whitney U	$3,442 \times 10^{-10}$
Intenzita nákladní automobily	Welchův t-test	$9,967 \times 10^{-05}$
Intenzita osobní automobily	Welchův t-test	$1,810 \times 10^{-35}$
Vyznamné plochy zeleňe	Welchův t-test	$2,421 \times 10^{-53}$

U všech zbylých datových sad, jak je vidět na obrázku 5.5 je průměrná teplota mimo oblasti vyšší. Dle výsledků testů prezentovaných v tabulce 5.7 zamítáme nulovou hypotézu, tudíž jsou teploty v oblasti jevu signifikantně nižší než mimo ně.

Po provedení statistických testů pro ověření navržených hypotéz tedy bylo zjištěno následující:

- Historické centrum a vilové čtvrti vykazují vyšší teploty, zatímco vnitřní centrum nevykazuje zvýšené teploty. Sídlíště a širší centrum také vykazují vyšší teploty, přičemž příměstská zástavba nevykazuje vyšší teploty.
- Zastavěné území a lomy vykazují vyšší teploty, zatímco vodní plochy, křoviny, orná půda, travní porosty a lesy vykazují nižší teploty.

- Industriální oblasti vykazují vyšší teploty, zatímco skládky a brownfieldy vykazují nižší teploty.
- Vodní toky, zelené střechy, intenzita nákladní dopravy, osobních automobilů a významné plochy zeleně vykazují nižší teploty.

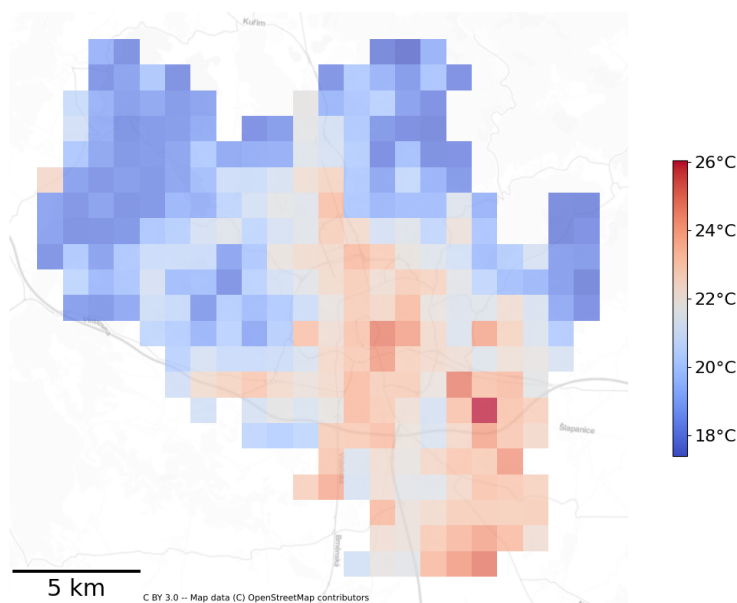
Shrnutí výsledků potvrzuje či vyvrací některé z navržených hypotéz:

- Industriální areály přispívají ke zvyšování teploty na svém území (hypotéza 1), zatímco skládky a brownfieldy ke zvyšování teploty oproti původnímu návrhu nepřispívají (hypotéza 1), (hypotéza 3).
- Využití půdy jako zelená plocha (les, křoviny, travní porost) (hypotéza 5) nebo jako vodní plocha přispívá ke snižování teploty na svém území (hypotéza 6).
- Využití půdy jako zastavěné území nebo lomy přispívá ke zvyšování teploty (hypotéza 7) a (hypotéza 4).
- Sídliště, historické centrum a vilové čtvrti přispívají ke zvyšování teploty na svém území, což částečně potvrzuje hypotézu 7, zatímco vnitřní centrum nevykazuje zvýšené teploty oproti původně navržené hypotéze.
- Významné plochy zeleně přispívají ke snižování teploty na svém území (hypotéza 8).
- Zelené střechy přispívají ke snižování teploty na své ploše (hypotéza 10).
- Ulice s vyšší intenzitou dopravy nezvyšují teplotu ve svém okolí oproti původně navržené hypotéze 9.

### 5.3 Postup při využití regrese a prostorové regrese

Při využití regrese bude zkoumán vliv jednotlivých datových vrstev na hodnoty povrchové teploty. Kromě knihoven již použitých v předchozí metodě tady bude využito funkcí z knihovny PySAL, která obsahuje užitečné funkce pro regresní analýzu.

Prvním krokem při této metodě je nutnost připravit a transformovat data, do formátu, ve kterém budou jednotlivé charakteristiky celé oblasti obsaženy v jednom GeoDataframu jako číselné hodnoty. Hlavním důvodem pro tuto transformaci je fakt, že datové sady ve formátu, v jakém jsou dostupné, obsahují vrstvu polygonů, případně multipolygonů, reprezentujících lokace daného jevu (např. lokace industriálních areálů, významných ploch zeleně a dalších). Řešením v této práci, podobně jako v [13], je vytvoření pravidelné mřížky polygonů ve tvaru čtverce o rozměrech 1000 m x 1000 m, přičemž tyto polygony budou obsahovat údaj o průměrné teplotě v jeho oblasti a procento plochy, které v dané buňce mřížky zaujímají jednotlivé jevy - nezávislé proměnné. Díky tomuto by metoda měla přinést důvěryhodnější výsledky o porovnání vlivu jednotlivých vrstev. Po vytvoření prázdných polygonů pokrývajících studovanou oblast je každému přiřazen identifikátor a pro všechny je nastaven souřadnicový systém S-JTSK. Poté je proveden výpočet průměrné teploty v každém z vygenerovaných polygonů mřížky jako vážený průměr zasahujících teplotních polygonů a jako váha je využita jejich plocha. Všechny datové sady, jejichž polygony označují lokace jevu



Obrázek 5.6: Mřížka teplotních polygonů na území města Brna

(např. industriální areály), jsou transformovány tak, aby se sloučily překrývající se polygony. U sad Využití půdy a Městská zástavba tato akce není nutná, jelikož jednotlivé druhy zástavby nebo využití jsou reprezentovány jako jeden polygon. Poté je vytvořen průnik všech sad s mřížkou teplot a je spočítána plocha daného jevu v dané buňce mřížky.

Z této plochy a plochy buňky je spočítáno, kolik procent plochy buňky je tvořeno daným jevem. Všechna tato data jsou sloučena do jednoho GeoDataframu obsahujícího všechny potřebné informace. Při těchto výpočtech může nastat, že některé oblasti nemají hodnotu – v tomto případě bude jako hodnota použita 0. Na tomto GeoDataframu byla použita metoda nejmenších čtverců (Ordinary Least Squares, OLS) 3.4 z knihovny pypsal<sup>1</sup>. Tato knihovna byla zvolena, protože narozdíl od jiných implementací OLS, lze zmínit například implementaci v knihovně Statsmodels<sup>2</sup> nebo Scikit-learn<sup>3</sup>, se knihovna pypsal zaměřuje cíleně na prostorovou analýzu. Kromě běžné OLS obsahuje také prostorové varianty OLS, z nichž je model prostorové závislosti využit dále. Pomocí funkce spreg.OLS() byl vytvořen model regrese, který zahrnuje závislou proměnnou 'Teplota' a nezávislé proměnné pro každou vrstvu nebo druh vrstvy.

### 5.3.1 Výsledky regresní analýzy

Tabulka 5.8: Výsledky klasické regrese

Vrstva	Koeficienty M1	Std. chyba M1	p-hodnota M1
Lomy	0,326	0,096832	0,000871
Historické centrum	0,0341	0,010878	0,001916
Industriální areály	0,0314	0,005448	$2,36661 \times 10^{-8}$
Vnitřní centrum	0,0216	0,007238	0,00315
Brownfieldy	0,0177	0,019957	0,376666
Širší centrum	0,0158	0,003272	$2 \times 10^{-6}$
Sídliště	0,0154	0,004775	0,001464
Vilové čtvrtě	0,0122	0,010553	0,247254
Orná půda	0,0117	0,004547	0,010539
Travní porost	0,0102	0,019561	0,60232
Příměstská zástavba	0,0066	0,001837	0,000364
Zastavěné území	0,0045	0,004067	0,270909
Skládky	-0,0042	0,008475	0,620899
Křoviny	-0,0157	0,006869	0,022794
Lesy	-0,0277	0,004358	$8,90969 \times 10^{-10}$
Významné plochy zeleně	-0,0387	0,030228	0,201818
Vodní plochy	-0,0457	0,01118	$5,8 \times 10^{-5}$
Vodní toky	-0,132	0,07001	0,060478
Zelené střechy	-1,4921	6,92146	0,829478

Důležitou metrikou modelu je tzv. koeficient determinace (R-squared) a jeho adjustovaná varianta (Adjusted R-squared), které ukazují, jak dobře model vysvětluje variabilitu závislé proměnné. Adjustovaný koeficient penalizuje modely s velkým počtem nezávislých proměnných, jelikož více těchto proměnných může uměle zvýšit klasický koeficient. Hodnota těchto metrik u tohoto modelu je 0,8273, respektive 0,8151 pro adjustovaný koeficient.

Největší vliv na ohřívání dle koeficientů tohoto modelu, zobrazeného v tabulce 5.8, mají vrstvy dolů a lomů, historické centrum, industriální areály a vnitřní centrum. Všechny tyto

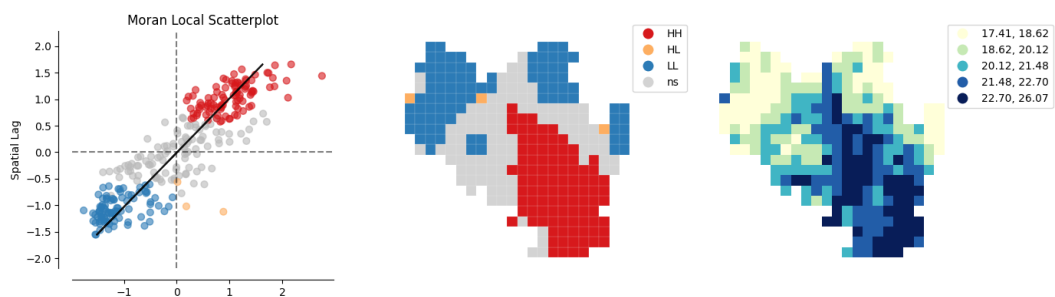
<sup>1</sup><https://pypsal.org/spreg/generated/spreg.OLS.html>

<sup>2</sup>[https://www.statsmodels.org/devel/generated/statsmodels.regression.linear\\_model.OLS.html](https://www.statsmodels.org/devel/generated/statsmodels.regression.linear_model.OLS.html)

<sup>3</sup>[https://scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LinearRegression.html](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html)

vrstvy mají p-hodnoty menší než je běžná hodnota 0,05, tudíž jsou tyto výsledky statisticky významné. Velmi nízký koeficient mají zelené střechy, ale jejich p-hodnota je velmi vysoká, tudíž tento výsledek není statisticky významný a vliv na teplotu nelze s jistotou posoudit. Podobně lze hovořit i o dalších vrstvách, jejichž p-hodnoty jsou vyšší než 0,05. Nejvyšší statisticky významný vliv na snižování teploty mají vodní toky a vodní plochy, menší vliv mají lesy a křoviny. Významné plochy zeleně mají vliv na snižování teploty, ale dle p-hodnoty se nejedná o statisticky významné výsledky.

Jak již bylo zmíněno v teoretické části, problémem pro běžnou metodu nejmenších čtverců může být prostorová autokorelace. Byla spočítána hodnota globální prostorové autokorelace, která je pro průměrnou teplotu 0,781 s p-hodnotou 0,001. Tyto hodnoty ukazují na vysokou míru autokorelace v teplotních datech, tudíž že oblasti s vyšší průměrnou teplotou se nacházejí poblíž jiných oblastí s vyšší teplotou a oblasti s nižší průměrnou teplotou se nacházejí poblíž dalších oblastí s nižší teplotou. P-hodnota ukazuje, že tato vyšší hodnota autokorelace je statisticky významná a je nepravděpodobné, že je tento výsledek náhodný.



Obrázek 5.7: Grafy autokorelace zobrazující míru autokorelace, její prostorové rozložení a statistickou významnost

Na prvním grafu na obrázku 5.7 lze pozorovat pozitivní trend v rámci globální korelace, jak je zmíněno výše.

Na druhém grafu lze vidět oblasti, kde je autokorelace významná. Červená barva – HH (High-High) označuje oblasti, kde jsou v blízkosti oblasti s vyšší teplotou, modrá barva – LL (Low-Low) označuje oblasti, kde jsou v blízkosti oblasti s nižší teplotou. HL a LH (High-Low a Low-High) popisují oblasti s negativní autokorelací, ty však ve studované oblasti nejsou významně zastoupeny. Šedá barva označuje oblasti, ve kterých není signifikantní žádná autokorelace.

Třetí graf ukazuje hodnoty statistické významnosti lokální korelace, nižší hodnoty znamenají vyšší statistickou významnost.

S autokorelací u závislé proměnné počítá tzv. Spatial lag model 3.5, jehož výsledky jsou v tabulce níže 5.9. Ukazateli kvality modelu jsou zde tzv. pseudo koeficient determinace a jeho prostorová varianta. Prostorová varianta vyjadřuje, jak dobře model vysvětluje závislou proměnnou bez započítání tzv. spatial lag komponentu, což znamená jak dobře vysvětlují závislou proměnnou pouze samotné nezávislé proměnné. Hodnota pro pseudo koeficient je 0,8846 a pro prostorovou variantu 0,8552, tedy vyšší než pro neprostorový model z tabulky 5.8. Nejvýznamnějšími vrstvami zvyšujícími okolní průměrnou teplotu, které jsou také statisticky významné, jsou ve výsledcích modelu prostorové závislosti, zobrazeného v tabulce 5.9, lomy, historické centrum, industriální areály, sídliště a vnitřní centrum. Zelené střechy mají opět nejnižší koeficient, avšak p-hodnota je stále velmi vysoká. Nejvýznamněj-



Tabulka 5.9: Výsledky prostorové regrese

Vrstva	Koeficienty M2	Std. chyba M2	p-hodnota M2
W_avg_temperature	0,4947	0,044385	$7,49296 \times 10^{-29}$
Lomy	0,304	0,076444	$7 \times 10^{-5}$
Historické centrum	0,0245	0,008629	0,00458
Industriální areály	0,0234	0,004345	$7,09236 \times 10^{-8}$
Sídliště	0,0128	0,003784	0,000694
Travní porosty	0,012	0,015443	0,436773
Vnitřní centrum	0,0119	0,005773	0,038587
Vilové čtvrtě	0,0097	0,008332	0,245583
Širší centrum	0,0076	0,002682	0,004538
Brownfields	0,004	0,015778	0,801057
Příměstská zástavba	0,004	0,001472	0,007236
Orná půda	0,0016	0,003651	0,658315
Zastavěné území	-0,0008	0,003241	0,8061
Křoviny	-0,0123	0,00543	0,023553
Skládky	-0,014	0,006736	0,037601
Les	-0,0208	0,003531	$3,58246 \times 10^{-9}$
Vodní plocha	-0,0329	0,008901	0,000218
Významné plochy zeleně	-0,0508	0,023868	0,033313
Vodní toky	-0,0978	0,05541	0,077419
Zelené střechy	-0,2594	5,46335	0,962127

šími vrstvami snižujícími okolní průměrnou teplotu a zároveň statisticky významnými jsou vrstvy vodních toků, zeleně, vodní plochy, lesa, křovin a skládek.

Tento model zohledňuje prostorovou závislost mezi jednotlivými pozorováními, což znamená, že vliv okolních oblastí na průměrnou teplotu je zahrnut do modelu. W\_avg\_temperature je koeficient prostorové závislosti, který ukazuje, jaký vliv má průměrná teplota okolních oblastí na průměrnou teplotu v dané oblasti. V tomto případě je koeficient W\_avg\_temperature 0,4955 a je statisticky významný, což znamená, že průměrná teplota v okolních oblastech má silný a pozitivní vliv na průměrnou teplotu v dané oblasti.

Následující tabulka 5.10 ukazuje porovnání koeficientů a jejich p-hodnot ve výsledcích obou modelů. Vrstvy jsou seřazené podle hodnot koeficientů z prostorového modelu, který by měl poskytovat důvěryhodnější výsledky kvůli vysoké míře prostorové autokorelace v závislé proměnné. Rozdíly nejsou příliš výrazné. Směr, kterým vrstvy ovlivňují teplotu, se změnil pouze u vrstvy zastavěné území, ale velmi málo a navíc jeho p-hodnota je velmi vysoká.

Tabulka 5.10: Porovnání výsledků běžné a prostorové regrese

Vrstva	Koeficienty M2	p-hodnota M2	Koeficienty M1	p-hodnota M1
Lomy	0,304	$7 \times 10^{-5}$	0,326	0,000871
Historické centrum	0,0245	0,00458	0,0341	0,001916
Industriální areály	0,0234	$7,09236 \times 10^{-8}$	0,0314	$2,36661 \times 10^{-8}$
Sídlště	0,0128	0,000694	0,0154	0,001464
Travní porosty	0,012	0,436773	0,0102	0,60232
Vnitřní centrum	0,0119	0,038587	0,0216	0,00315
Vilové čtvrtě	0,0097	0,245583	0,0122	0,247254
Širší centrum	0,0076	0,004538	0,0158	$2 \times 10^{-6}$
Brownfieldy	0,004	0,801057	0,0177	0,376666
Příměstská zástavba	0,004	0,007236	0,0066	0,000364
Orná půda	0,0016	0,658315	0,0117	0,010539
Zastavěné území	-0,0008	0,8061	0,0045	0,270909
Křoviny	-0,0123	0,023553	-0,0157	0,022794
Skládky	-0,014	0,037601	-0,0042	0,620899
Lesy	-0,0208	$3,58246 \times 10^{-9}$	-0,0277	$8,90969 \times 10^{-10}$
Vodní plocha	-0,0329	0,000218	-0,0457	$5,8 \times 10^{-5}$
Významné plochy zeleně	-0,0508	0,033313	-0,0387	0,201818
Vodní toky	-0,0978	0,077419	-0,132	0,060478
Zelené střechy	-0,2594	0,962127	-1,4921	0,829478

Na základě srovnání výsledků prostorové regrese (M1) a obyčejné regrese (M2) v tabulce 5.10 lze učinit následující závěry týkající se vlivu využití půdy na teplotu:

- Zelené plochy (lesy, keře, trávníky) a vodní plochy přispívají ke snižování teploty v jejich okolí (hypotéza 5).
- Zastavěné území (hypotéza 7) a lomy přispívají ke zvyšování teploty (hypotéza 4).
- Sídlště, historická centra, vilové čtvrti a vnitřní centrum přispívají ke zvyšování teploty v jejich okolí (hypotéza 7).
- Významné zelené plochy přispívají ke snižování teploty v jejich okolí (hypotéza 8).
- Industriální areály mírně přispívají ke zvyšování teploty (hypotéza 1), skládky a brownfieldy ve srovnání s původní hypotézou nemají významný vliv (hypotéza 3 a 2)
- Nelze se statistickou významností říci, že zelené střechy přispívají ke snižování teploty na jejich povrchu v porovnání s původní hypotézou 10

Většina navržených hypotéz byla potvrzena, avšak s různým stupněm významnosti. Významné zvýšení teploty bylo zjištěno v průmyslových oblastech, historických centrech, bytových oblastech a širším centru města. Doly a lomy měly také významný vliv na zvýšení teploty. Vnitřní město a příměstská zástavba vykazovaly mírně významný vliv na zvýšení teploty. Naopak orná půda prokázala mírně významný vliv na snižování teploty pouze v modelu M2.

Některé hypotézy nebyly potvrzeny. Například zelené střechy, vilové čtvrti, zastavěné území, travní porost a brownfieldy neukázaly významný vliv na teplotu ani v jednom modelu. Zelené střechy měly podle obou modelů mít významný vliv na snižování teploty, avšak tento fakt nebyl podpořen se statistickou významností. Keře prokázaly mírně významný vliv na snižování teploty v obou modelech, stejně tak vodní plochy a toky vykazovaly mírně významný vliv na snižování teploty.

## 5.4 Vybrané potvrzené teorie podrobněji

Na základě předchozích výsledků byly vybrány vrstvy, na kterých lze identifikovat konkrétní jednotlivé prvky (jako např. jednotlivé parky) a prozkoumat jejich vliv.

### 5.4.1 Lomy

Vrstva lomů na území Brna, zobrazena na obrázku 5.8, obsahuje dvě oblasti, ve kterých se lomy vyskytují. Jednou z nich je lom Kalcit a bývalý lesní lom blízko části Brno-Líšeň, kde je průměrná teplota mezi 20 a 21 stupni. Druhý výskyt lomů je v oblasti pískovny Černovice. Tyto oblasti mají výrazně vyšší teplotu 25 až 27 stupňů, přičemž průměrná teplota v celé oblasti Brna je asi 22,3 stupňů. Lomy blízko Líšně nemají vliv na oteplování, což může být způsobené i přítomností lesa v jejich okolí, zatímco oblast v Černovicích je výrazně teplejší než průměr v celém městě. Ačkoli mají některé lomy vyšší teplotu než je průměr v Brně, jejich plocha představuje pouze 0,06 % z celé plochy města, jak lze vidět v legendě na obrázku 4.3.



Obrázek 5.8: Lomy na území Brna

### 5.4.2 Významné zelené plochy

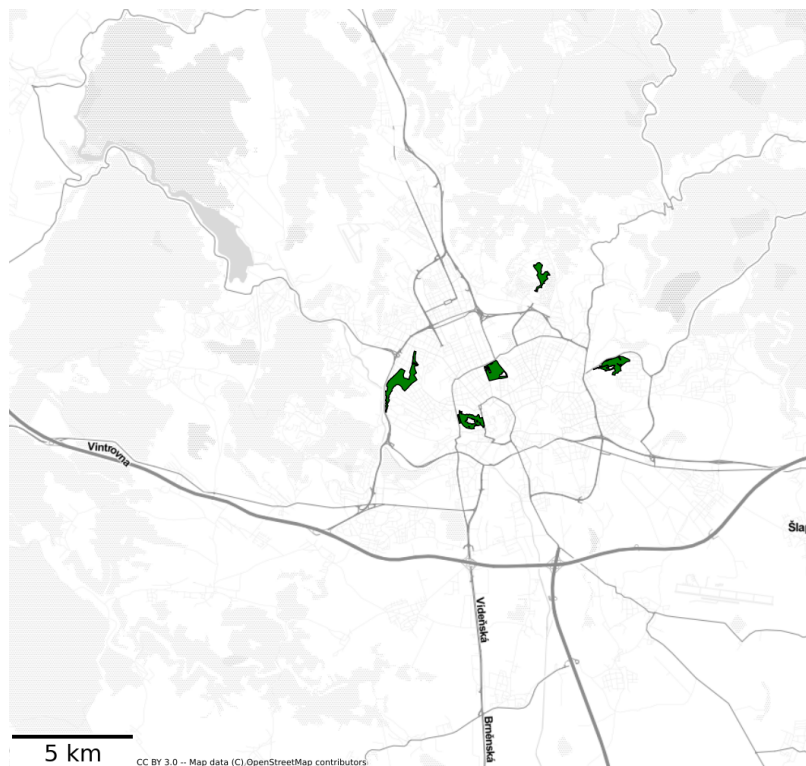
Největšími zástupci datové sady významných ploch zeleně, prezentovány v tabulce 5.11 a vizualizovány na obrázku 5.9, jsou velké parky v centru, případně vnějším centru města. Jak je vidět při porovnání s nejchladnějšími zelenými plochami, prezentovanými v tabulce 5.12, jejich průměrné teploty jsou blízko těmto plochám. Wilsonův les je dokonce druhou nejchladnější plochou. Všechny největší plochy mají také nižší teplotu, než je průměrná teplota v celém městě. Celkově lze pozorovat, že již v blízkém okolí 50 metrů okolo zelených ploch teplota stoupá a tento trend pokračuje i ve větším okolí 200 metrů okolo zelených ploch. Vliv zelených ploch na lokální teplotu je tedy značný. Protože okolo nich je většinou zastavěná oblast v centru města, tvoří tedy chladné ostrovy.

Tabulka 5.11: 5 největších zelených ploch

Název lokality	Průměrná teplota ve °C	Plocha v $m^2$	Průměrná teplota v okolí 50 m ve °C	Průměrná teplota v okolí 200 m ve °C
Wilsonův les	19,8	788 945	19,8	20,6
Akátky	20,7	442 505	21,4	21,4
Lužánky	21,0	424 064	22,3	22,4
Špilberk	21,5	400 331	21,8	21,9
Čertova rokle	20,5	272 888	21,4	22,3

Tabulka 5.12: 5 nejchladnějších zelených ploch

Název lokality	Průměrná teplota ve °C	Plocha v $m^2$	Průměrná teplota v okolí 50 m ve °C	Průměrná teplota v okolí 200 m ve °C
Park u pomníku	19,5	5 203	20,5	20,5
Wilsonův les	19,8	788 945	19,8	20,6
Ondrova - pod ZOO	20,0	6 149	20,0	20,7
Anthropos	20,2	77 231	20,9	21,6
Návrší Svobody - Potocká	20,4	9 158	20,5	20,9



Obrázek 5.9: 5 největších zelených ploch z datasetu Významné zelené plochy

### 5.4.3 Industriální areály

Největší areály, prezentovány v tabulce 5.14, mají výrazně vyšší průměrnou teplotu na svém území než je celková průměrná teplota celého města. Nejvyšší teplota z 5 největších je v areálu Královopolské a.s., která je představitelem strojího průmyslu, stejně jako areál ZETOR a.s. Brněnský autodrom má výrazně nižší teplotu než ostatní velké areály a nejteplejší areály. Jako nejteplejší areály se ukazují strojírenské podniky nebo areál brněnských tepláren. Při porovnání teplot v blízkém a vzdálenějším okolí, tedy 50 a 200 metrů, lze pozorovat podobný efekt jako u zelených ploch, ale v opačném směru. Teploty již v blízkém okolí jsou až o jednotky stupňů nižší než přímo v jednotlivých areálech a dále se snižují se zvětšující se vzdáleností od areálů.

Tabulka 5.13: 5 největších areálů

Název areálu	Průměrná teplota ve °C	Plocha v $m^2$	Průměrná teplota v okolí 50 m ve °C	Průměrná teplota v okolí 200 m ve °C
Letiště Brno - Tuřany	25,1	4 661 227	23,0	23,0
ZETOR, a.s.	24,6	2 355 601	22,8	22,7
Autodrom Brno	21,5	1 894 506	21,2	21,0
Královopolská a.s. - původní rozsah areálu	25,9	876 726	23,9	23,5
Depo kolejových vozi- del (DKV)	23,4	650 848	23,1	22,6

Tabulka 5.14: 5 nejteplejších areálů

Název areálu	Průměrná teplota ve °C	Plocha v $m^2$	Průměrná teplota v okolí 50 m ve °C	Průměrná teplota v okolí 200 m ve °C
Honeywell, spol. s r.o. - Brno o.z.	29,0	264 021	28,7	28,2
Izolas, spol. s r.o.	28,3	16 290	26,9	25,7
OLAER, s.r.o.	27,5	1 480	26,4	25,3
Chepos Engineering	27,3	19 159	26,8	26,6
Královopolská a.s. - současný provoz Chemie I	26,8	56 912	25,9	25,6

## Kapitola 6

# Závěr

Cílem práce bylo s využitím dostupných datových sad analyzovat jejich vztah s teplotou povrchu na území města Brna. Pro vypracování bylo nutné nastudovat formáty geografických dat a způsoby práce s nimi. Dále bylo potřeba prozkoumat možnosti využití statistického testování hypotéz a statistických testů. Bylo také nutné prozkoumat možnost využití regresní analýzy a její limitace při práci s prostorovými daty. Řešením těchto limitací je využití prostorového regresního modelu. Dalším krokem bylo prezentování vybraných datových sad a navržení hypotéz vzhledem k jejich potenciálnímu vztahu ke zvyšování nebo snižování teploty povrchu. Nastudované metody byly využity k ověření těchto hypotéz. Některé potvrzené hypotézy byly na konci práce rozebrány podrobněji.

Výsledkem práce jsou metody pro statistické testování a pro regresní analýzu implementované v programovacím jazyce Python, výsledky získané pomocí těchto metod a jejich následná interpretace vzhledem k původně navrženým hypotézám. Analýza se zabývá konkrétními daty zaměřenými na oblast města Brna, ale v případě dostupnosti dat pro jiné oblasti je možné využít implementované metody pro analýzu jiných oblastí.

Limitací práce je jistě fakt, že teplotní data jsou pořizena v jeden konkrétní moment, a tudíž mohla být ovlivněna např. aktuálními povětrnostními podmínkami. Závěry na základě analýzy této jedné datové sady mohou být tímto faktem zkreslené. Dostupnost více datových sad pro různé časy dne by umožnila podrobnější časovou analýzu dat. Dostupnost, v případě historických teplotních dat a dalších vrstev by mohla být provedena analýza dlouhodobých trendů, která by pomohla pochopit, jak se teplota vyvíjela v průběhu času. To by mohlo pomoci identifikovat oblasti, které zažily významné teplotní změny a možné faktory přispívající k těmto změnám. Tato fakta by bylo možné využít při plánování výstavby nových částí města nebo přestavbě těch stávajících.



# Literatura

- [1] ANSELIN, L. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers, 1988. ISBN 978-0-89838-575-7.
- [2] ARRIBAS BEL, D. a WOLF, L. *Geographic Data Science with Python*. 2021. Dostupné z: <https://geographicdata.science/book/intro.html>.
- [3] BOGNER, J., ABDELRAFIE AHMED, M., DIAZ, C., FAAIJ, A., GAO, Q. et al. Waste management. In: *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC, 2007. Dostupné z: [https://archive.ipcc.ch/publications\\_and\\_data/ar4/wg3/en/ch10.html](https://archive.ipcc.ch/publications_and_data/ar4/wg3/en/ch10.html).
- [4] BONAN, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*. American Association for the Advancement of Science. 2008, sv. 320, č. 5882, s. 1444–1449.
- [5] BUTLER, H., DALY, M., DOYLE, A., GILLIES, S., HAGEN, S. et al. *The GeoJSON Format* [Internet Requests for Comments]. RFC 7946. RFC Editor, August 2016. Dostupné z: <https://www.rfc-editor.org/rfc/rfc7946>.
- [6] CHANG, C.-R., LI, M.-H. a CHANG, S.-D. A preliminary study on the local cool-island intensity of Taipei city parks. *Landscape and urban planning*. Elsevier. 2007, sv. 80, č. 4, s. 386–395.
- [7] DRAPER, N. R. a SMITH, H. *Applied Regression Analysis*. 3. vyd. John Wiley & Sons, 1998.
- [8] DU, H., SONG, X., JIANG, H., KAN, Z., WANG, Z. et al. Research on the cooling island effects of water body: A case study of Shanghai, China. *Ecological indicators*. Elsevier. 2016, sv. 67, s. 31–38.
- [9] ESRI. *ESRI Shapefile Technical Description*. 1998. Dostupné z: <https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf>.
- [10] FIELD, A. *Discovering Statistics Using IBM SPSS Statistics*. SAGE Publications, 2013. Introducing statistical methods. ISBN 9781446249178.
- [11] FINDA, J. *Kartografická zobrazení [online]*. 2006 [cit. 2023-03-05]. Diplomová práce. Masarykova univerzita, Přírodovědecká fakulta, Brno. SUPERVISOR : Josef Janyška. Dostupné z: <https://is.muni.cz/th/dv9gw/>.

- [12] GLINER, J., MORGAN, G. a LEECH, N. *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis*. únor 2000. ISBN 9781410605337.
- [13] HART, M. A. a SAILOR, D. J. Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theoretical and applied climatology*. Springer. 2009, sv. 95, s. 397–406.
- [14] HEYWOOD, D., CORNELIUS, S. a CARVER, S. *An Introduction to Geographical Information Systems*. Pearson Prentice Hall, 2006. Pearson educación. ISBN 9780131293175.
- [15] MAHMOOD, R., PIELKE SR, R. A., HUBBARD, K. G., NIYOGI, D., DIRMEYER, P. A. et al. Land cover changes and their biogeophysical effects on climate. *International journal of climatology*. Wiley Online Library. 2014, sv. 34, č. 4, s. 929–953.
- [16] MORAN, P. A. Notes on continuous stochastic phenomena. *Biometrika*. JSTOR. 1950, sv. 37, 1/2, s. 17–23.
- [17] OKE, T. R. The energetic basis of the urban heat island. *Quarterly journal of the royal meteorological society*. John Wiley & Sons, Ltd Bracknell. 1982, sv. 108, č. 455, s. 1–24.
- [18] PATAKI, D. E., CARREIRO, M. M., CHERRIER, J., GRULKE, N. E., JENNINGS, V. et al. Coupling biogeochemical cycles in urban environments: ecosystem services, green solutions, and misconceptions. *Frontiers in Ecology and the Environment*. Wiley Online Library. 2011, sv. 9, č. 1, s. 27–36.
- [19] RATLIFF, E. *Google Maps Is Changing the Way We See the World* [online]. Brno: wired.com, červen 2007 [cit. 2023-03-05]. Dostupné z: <https://www.wired.com/2007/06/ff-maps/>.
- [20] SANDVIK, B. Using KML for Thematic Mapping. *AGU Fall Meeting Abstracts*. Prosinec 2008. Dostupné z: [https://thematicmapping.org/downloads/Using\\_KML\\_for\\_Thematic\\_Mapping.pdf](https://thematicmapping.org/downloads/Using_KML_for_Thematic_Mapping.pdf).
- [21] SANTAMOURIS, M., PAPANIKOLAOU, N., LIVADA, I., KORONAKIS, I., GEORGAKIS, C. et al. On the impact of urban climate on the energy consumption of buildings. *Solar energy*. Elsevier. 2001, sv. 70, č. 3, s. 201–216.
- [22] SANTAMOURIS, M. Cooling the cities—a review of reflective and green roof mitigation technologies to fight heat island and improve comfort in urban environments. *Solar energy*. Elsevier. 2014, sv. 103, s. 682–703.
- [23] ULPIANI, G. On the linkage between urban heat island and urban pollution island: Three-decade literature review towards a conceptual framework. *Science of the Total Environment*. Elsevier. 2021, sv. 751, s. 141727.
- [24] ZHOU, B., RYBSKI, D. a KROPP, J. P. The role of city size and urban form in the surface urban heat island. *Scientific reports*. Springer. 2017, sv. 7, č. 1, s. 1–9.