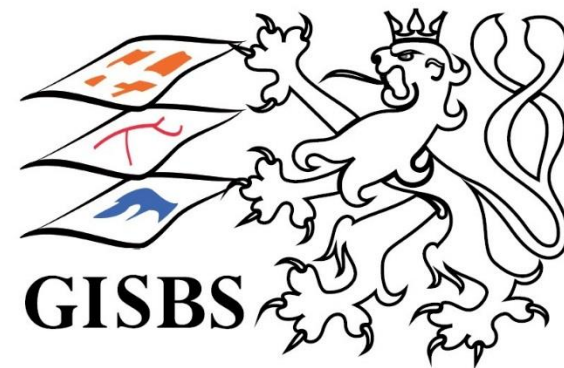




MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



Analýzy kriminality v rámci projektu „Geoinformatika jako nástroj pro podporu integrované činnosti bezpečnostních a záchranných složek státu“

Doc. Dr. Ing. Jiří Horák, VŠB-TU Ostrava

Doc. Dr. Ing. Bronislava Horáková, Intergraph CS s.r.o.

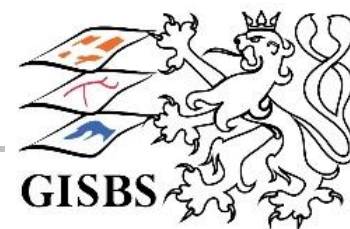
Program „Bezpečnostní výzkum pro potřeby státu v letech 2010 až 2015“

č.j. MV-32046-58/VZ-2012

 INTERGRAPH®



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



Dodavatel

Intergraph CS s.r.o



Subdodavatelé

Vysoká škola báňská - Technická univerzita Ostrava

☐ Institut geoinformatiky

Masarykova univerzita Brno

☐ Geografický ústav

Gefos, a.s.



Odborní gestoři projektu:

Policejní prezidium ČR, oddělení krizového řízení

V rámci Policie České republiky má odpovědnost za projekt pracovní skupina k zabezpečení úkolů PČR v tomto projektu.

Ministerstvo vnitra ČR, odbor eGovernmentu

Ministerstvo vnitra ČR, odbor prevence kriminality



Primární cíle projektu:

- ☐ metodicky navrhnout a prakticky odzkoušet geoinformační technologie ve specifickém prostředí PČR
- ☐ zajistit ze strany MV a PČR podporu koordinované činnosti bezpečnostních a záchranných složek státu sdílením standardizovaných postupů, prostorových dat, doprovodných informací a výstupů z analýz ve společném prostředí

Požadované výsledky:

- ☐ Pilotní implementace vhodných GIT ve specifickém prostředí PČR
- ☐ Získané poznatky přenést do podoby certifikované metodiky

Tematické oblasti:

- ☐ Krizové řízení a výkon služby
- ☐ Prevence a kriminalita

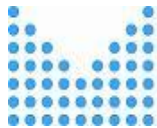
Technologie Intergraph nasazené v projektu:

- ☐ GSP Server Professional
- ☐ GeoMedia Smart Client Professional
- ☐ Analytické moduly GeoMedia Grid, GeoMedia 3D, Imagine Professional
- ☐ GeoMedia Professional a GeoMedia Fusion
- ☐ Intergraph Planning and Response Professional



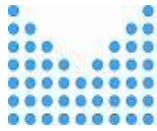
Analýzy kriminality

- IACA (Mezinárodní asociace kriminálních analytiků) vymezuje základní typy analýzy kriminality na základě povahy a zdroje dat, použitých technik, výsledků, frekvence a účelu. Jedno z možných členění:
 - informační analýza kriminality (crime intelligence analysis) – sběr informací a jejich analýza, kontext a motivy kriminality, recidivisté, opakované oběti, organizované skupiny
 - taktická analýza kriminality (tactical crime analysis) - krátkodobý vývoj kriminality za účelem efektivního rozmístění SaP policie, analýza opakovaných incidentů, analýzy vzorů chování a propojování pachatelů k minulým deliktům
 - strategická analýza kriminality (strategic crime analysis) - dlouhodobé strategie a prevence kriminality; analýzy trendu, analýzy hotspots a analýzy problémů
 - administrativní analýza kriminality (administrative crime analysis) - statistické hodnocení kriminality, výkazy, hodnocení činnosti, žádosti o granty, média



Etapy projektu GISBS

- Analýza stavu předmětné problematiky a implementace pilotního řešení
- Datová fúze zájmových vnitřních zdrojů dat
- Datová fúze zájmových vnějších zdrojů dat
- Vypracování certifikované metodiky



Analýza a návrh sady analytických funkcí pro analýzy kriminality a koncepce jejich využití

- prozkoumána řada metod
- popisy funkcionality potřebné pro provádění analýz
- příprava doporučení



Explorační a prostorová explorační analýza

- **výpočet základních neprostorových charakteristik:** aritmetický průměr, medián, modus, dolní a horní kvartil, rozptyl, směrodatná odchylka, variační koeficient, minimum, maximum, koeficient šikmosti a koeficient špičatosti
- **výpočet základních prostorových charakteristik:** průměrný střed, mediánový střed (podle mediánu souřadnic), směrodatná vzdálenost, elipsa standardizované odchylky, konvexní obálka. Vše s možností zadat váhy pro jednotlivé objekty
- **sloupcový graf** (vhodné nastavování, interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji),
- **výsečový graf** (vhodné nastavování, interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji),
- **krabicový graf** (interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji)
- **histogram** (volitelné šířky tříd, interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji),
- **sloupcový a kruhový kartodiagram** (vhodné nastavování, interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji)
- **kartogram s možností klasifikace hodnot** podle kvantilů (kvantilová mapa), podle směrodatné odchylky, podle metody přirozených zlomů, rovnoměrné členění, uživatelské členění (klasifikační mapa). Vhodné nastavování, interaktivita kurzorovým výběrem, propojení s ostatními grafickými nástroji.

Matematické postupy, ukázky řešení

Postup:

1) vypočtou se souřadnice průměrného středu \bar{X} , \bar{Y} .

2) pro každý bod se vypočtou transformované souřadnice

$$x'_i = x_i - \bar{X} \quad y'_i = y_i - \bar{Y}$$

3) vypočte se úhel rotace

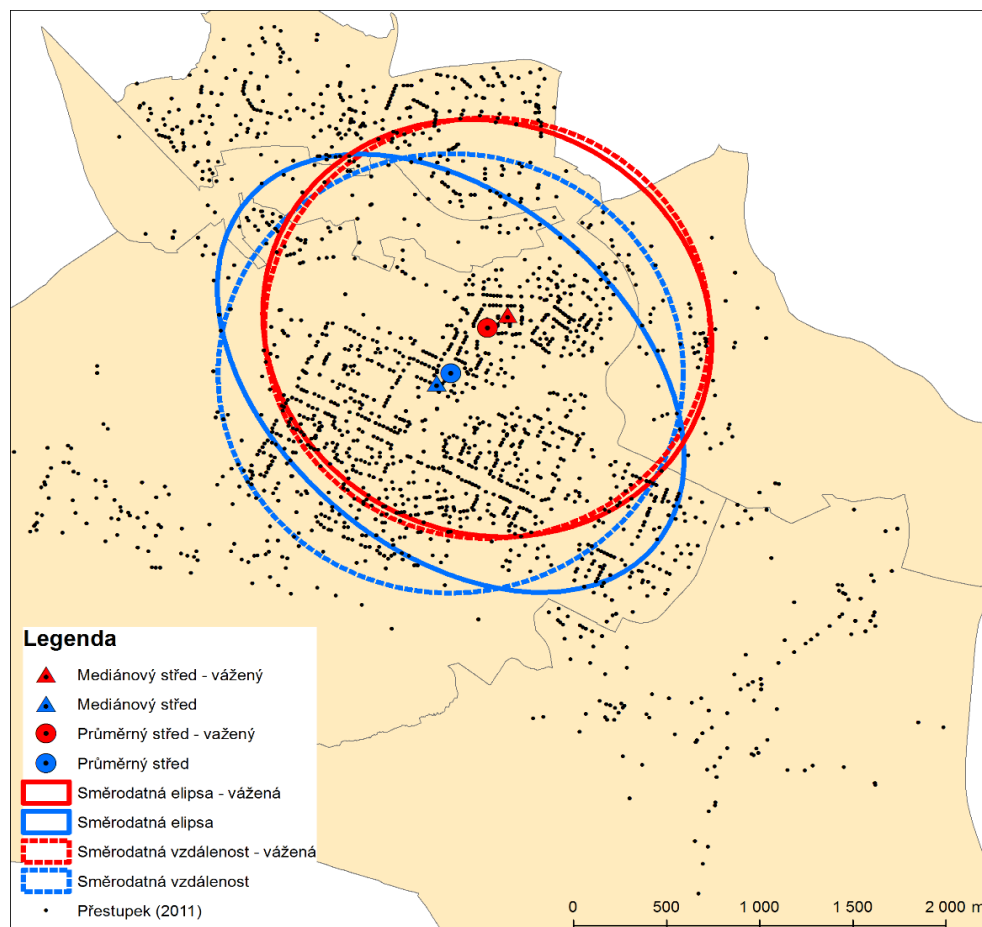
$$\tan \theta = \frac{\left(\sum_{i=1}^n x_i'^2 - \sum_{i=1}^n y_i'^2 \right) + \sqrt{\left(\sum_{i=1}^n x_i'^2 - \sum_{i=1}^n y_i'^2 \right)^2 + 4 \left(\sum_{i=1}^n x_i' \sum_{i=1}^n y_i' \right)^2}}{2 \sum_{i=1}^n x_i' \sum_{i=1}^n y_i'}$$

4) vypočtou se délky poloos elipsy

$$l_a = \sqrt{\frac{\sum_{i=1}^n (x'_i \cos \theta - y'_i \sin \theta)^2}{n}} \quad l_b = \sqrt{\frac{\sum_{i=1}^n (x'_i \sin \theta + y'_i \cos \theta)^2}{n}}$$

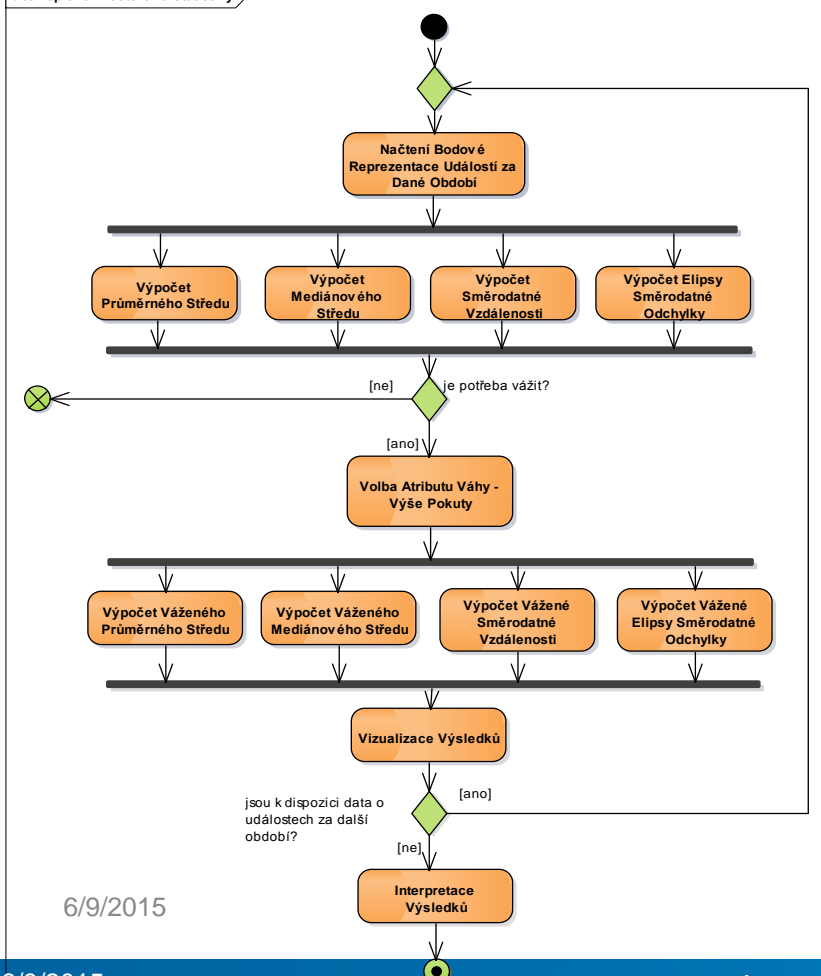
V CrimeStat a Smith et al. (2011) se uvádí mírně odlišný postup výpočtu směrodatných odchylek ve směru obou os:

$$SD_x = \sqrt{\frac{2 \sum_{i=1}^n ((x_i - \bar{x}) \cos \theta - (y_i - \bar{y}) \sin \theta)^2}{n-2}} \quad SD_y = \sqrt{\frac{2 \sum_{i=1}^n ((x_i - \bar{x}) \sin \theta + (y_i - \bar{y}) \cos \theta)^2}{n-2}}$$

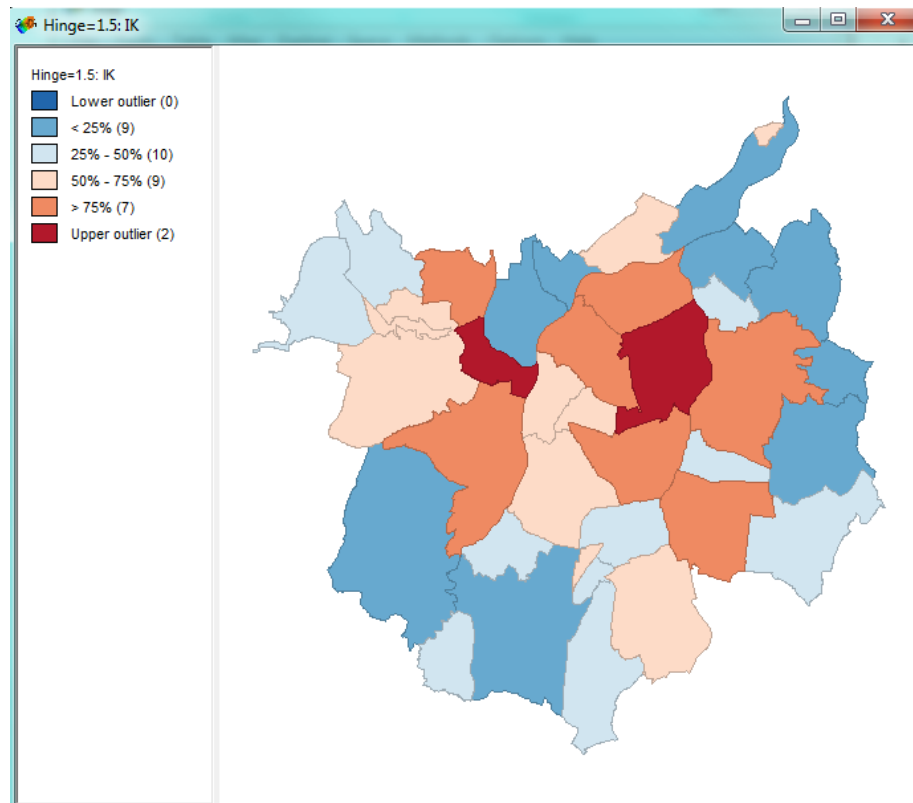


Diagramy využití, ukázky řešení

act Popisné Prostorové Statistiky



6/9/2015

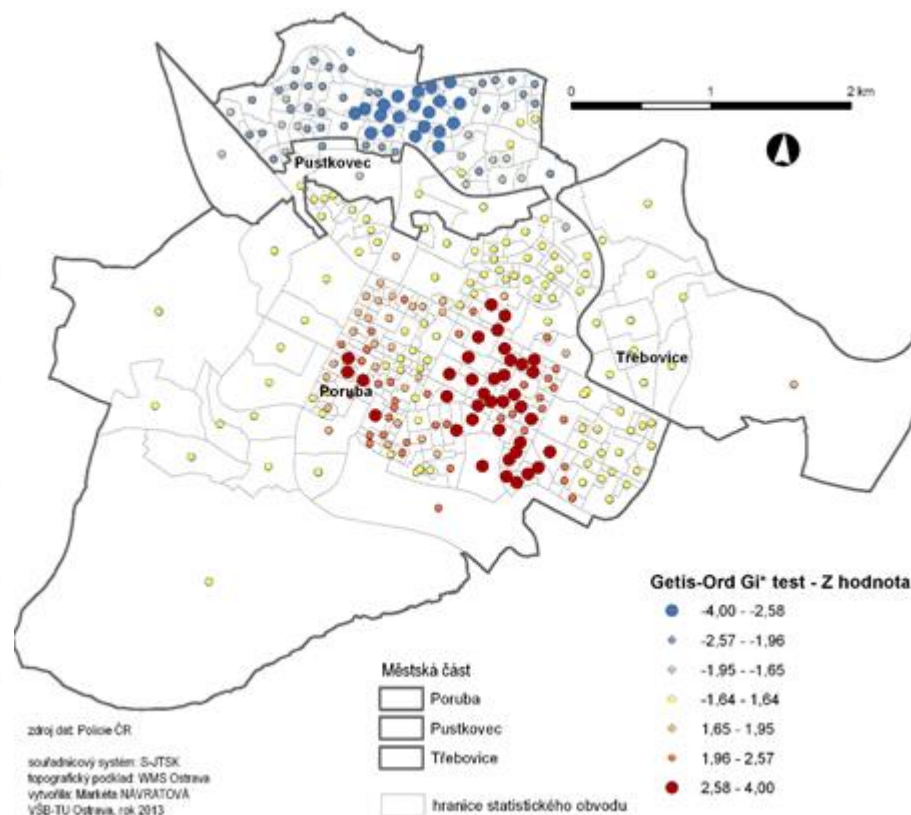
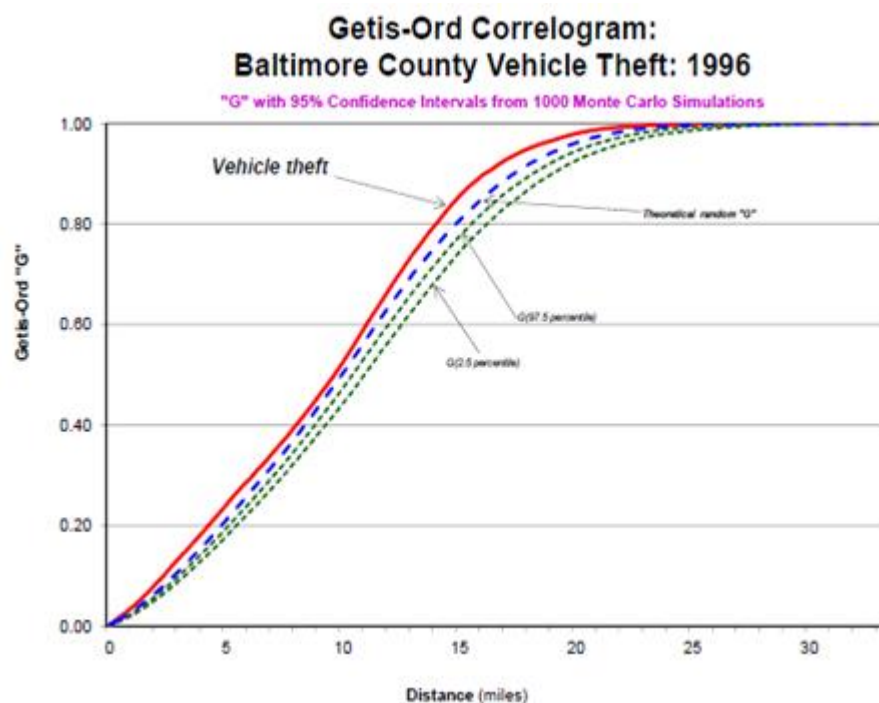




Hodnocení autokorelace pro polygony (čtverce rastru)

- **Moranovo / kritérium** (globální i lokální verze, možnost volit různé prostorové schéma vah, využití permutační Monte Carlo simulace pro ověření významnosti, zobrazování Moranova bodového grafu, mapové zobrazení lokální statistiky),
- **Gearyho C kritérium** (možnost volit různé prostorové schéma vah, využití permutační Monte Carlo simulace pro ověření významnosti, zobrazování korelogramu)
- **Getis-Ord G a lokální G_i^*** (možnost volit různé prostorové schéma vah, využití permutační Monte Carlo simulace pro ověření významnosti, zobrazování korelogramu pro globální funkci, mapové zobrazení lokální statistiky)
- Vlastnost volit různé prostorové schéma vah by měla zahrnovat varianty: topologické sousedství (královna, věž), prahová vzdálenost s přednastavenou minimální vzdáleností, k nejbližších sousedů.

Vývojové diagramy postupů zpracování a vyhodnocení





Prostorové interferenční testy

■ kvadrantové analýzy

- testování VMR indexu disperze nebo X^2 testu.
- Metoda je vhodná především kvůli jejich jednoduchosti a možnosti multivariačního zpracování (více atributů, kombinace se socioekonomickými daty apod.) a časoprostorového zpracování.

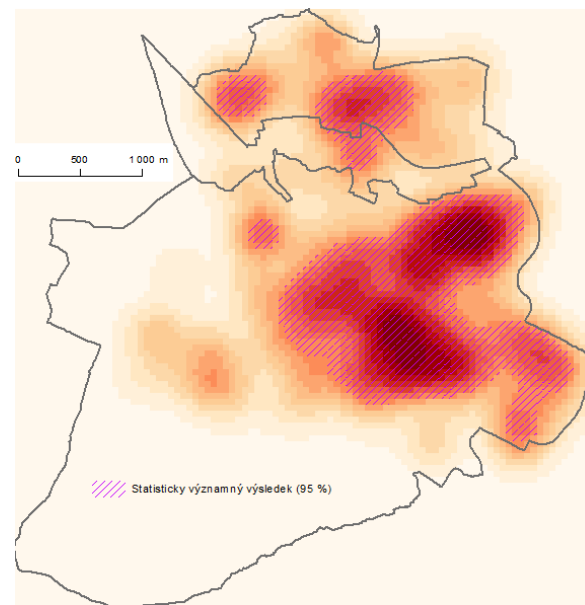
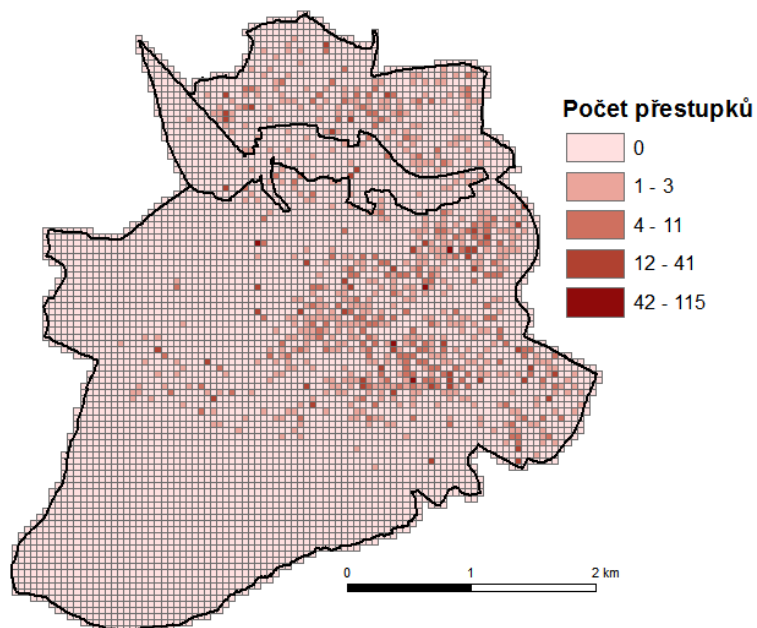
- **L funkce** (vhodná transformace Ripleyovy K funkce) - umožňuje hodnocení situace v závislosti na měřítku pozorování a rovněž je vhodná i pro adjustaci dat a relativizaci vůči socioekonomickým nositelům



Hot spots analýzy

- **kvadrantové metody** - vhodná především kvůli jejich jednoduchosti a možnosti multivariačního zpracování (více atributů, kombinace se socioekonomickými daty apod.) a časoprostorového zpracování)
- **jádrově založené metody** vhodná varianta adaptivního jádrového odhadu, typy jádrových funkcí – normální Gaussova, homogenní/uniformní, kvartická, kvadratická - Epanechnikov).
- **jádrové odhady na liniích**
- Vhodné doplnění:
 - Nearest Neighbour Hierarchical Clustering
 - další multivariační shlukovací techniky

Hot spots analýzy





Analýza potřeb v souvislosti s požadavky analytického zpracování a situací v datových zdrojích

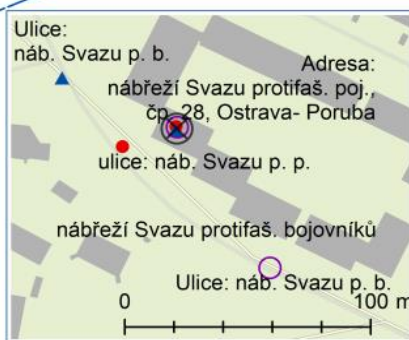
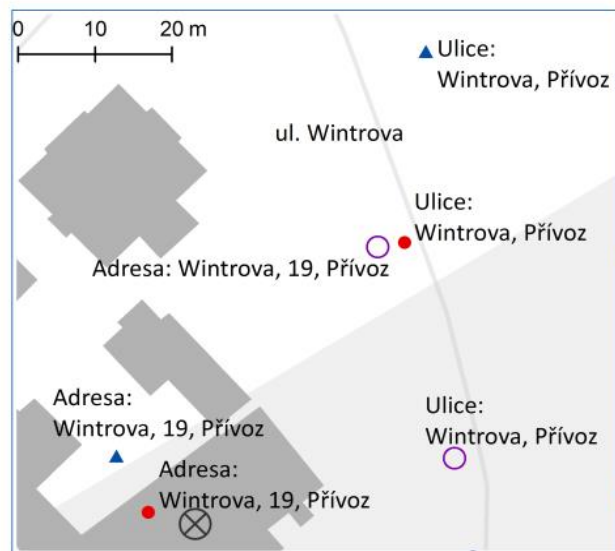
- analýza dodaných podkladů, dat
- jejich rozbor pro základní charakteristiku situace v interních a externích datových zdrojích,
- zpracování vzorků dodaných dat z nejméně 2 lokalit – Kolín, Ostrava
- Zjišťovaly se parametry dat, provádělo se hodnocení struktury, sémantiky, naplnění, obsahu a problémů, které se vyskytují v jednotlivých datových sadách:
 - Data PČR
 - Data Městské policie Ostrava
 - Data Městské policie Kolín
- popisuje se nutnost a možnost harmonizace dat a jejich integrace (společné číselníky, vazby)



Zhodnocení metod geokódování

- Návrh na předzpracování dat a jejich geokódování
 - Charakteristika geokódování
 - Geokódování pomocí API mapy.cz
 - Geokódování pomocí Google Maps Geocoding API
 - Geokódování pomocí projektu Nominatim (součást OpenStreetMap)
 - Geokódování pomocí MapQuest
 - Geokódování pomocí Bing Maps
 - Databázové geokódování
 - Vlastní aplikace Geocoder (využívá fulltextové vyhledávání)
 - Popis přípravy geokódování
 - Popis harmonizace zápisu lokalizačních údajů
- Aplikace předzpracování dat a geokódování na evidenci přestupků Městské policie Ostrava
 - Z celkového počtu 117921 záznamů chybí souřadnice 780 záznamů (0,7%).
 - kvalitní lokalizace 83348 záznamů
 - dalších 34573 (29%) určeny aproximativně
 - Z toho 26945 (23%) určené pouze na střed ulice.

VÝSLEDKY GEOKÓDOVÁNÍ POMOCÍ MAPOVÝCH VYHLEDÁVAČŮ



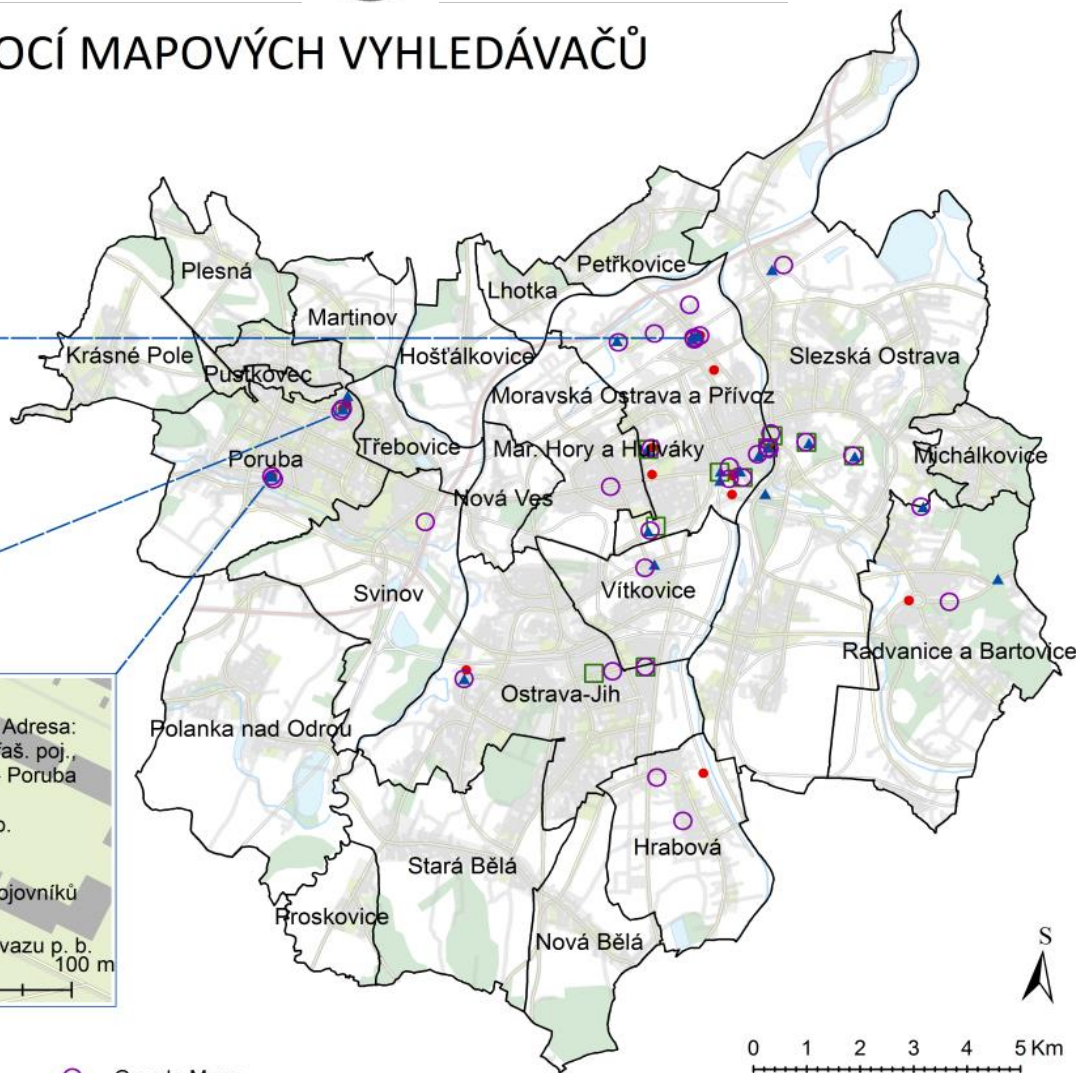
Bc. Pavel Vícha

VŠB-TUO, Ostrava 2015

Souřadnicový systém: S-JTSK

Zdroj dat: Městská policie Ostrava

Zdroj podkladových dat: Magistrát města Ostravy



- Google Maps
- MapQuest
- Bing Maps
- Mapy.cz



Skutečná poloha



Hranice měst. částí



Ulice

Dálnice

Silnice

Železnice

Řeky

Budovy

Vodní plochy

Les

Zeleň



Metodika harmonizace, agregace a anonymizace dat ETŘ

- Cíle - odstranění některých problémů, zvýšení kvality evidence, nastavení vhodných mechanismů pro automatizaci některých činností (např. generování statistických sestav)
- určena pro programátory, editory a příp. analytiky k zajištění jednotného postupu předzpracování dat a jejich interpretace, kontrole zapsaných dat, vytváření harmonizovaných údajů pro analytické potřeby, zejména hodnocení stavu a vývoje kriminality a návrhy preventivních opatření
- základní využití metodiky bude na krajské úrovni (krajské ředitelství PČR), s doporučenou frekvencí 1 x denně. Implementace co nejvíce automatizovaně.
- 3 základní části – harmonizace dat, agregace dat a anonymizace dat



Harmonizace dat ETŘ

- **Kontrola číselníků**
 - Ověření referenční integrity (5 kontrol, např. TRIDA_KOD obsahuje pouze kódy uvedené v číselníku TSK)
 - Kontrola obsahu (3 kontroly, např. kontrola, zda událost řeší pouze útvar správného typu)
- **Kontrola konzistence časových údajů**
 - Kontrola daleké minulosti nebo naopak budoucnosti (4 kontroly, např. rozdíl mezi Dat_DO a Dat_OD není větší než 10 let)
 - Zajištění logické návaznosti časů (2 kontroly, např. DAT_OD je menší nebo rovno DAT_DO)
- **Kontrola konzistence souřadnic**
 - Automatizovaná kontrola nekonzistentních zadaných souřadnic (4 kontroly, např. mimo rozsah ČR, místo spáchání služebna PČR)
- **Výpočet odvozených atributů**
 - Výpočet referenčního času spáchání události (vč. nepřesných údajů, ošetření zvláštních TSK)
- **Doplnění evidence objektů**
- **Doplnění klasifikace události** - doplnění kódu nadtřídy a kódu kategorie TSK, pokud je znám pouze kód třídy, případně zpřesnění klasifikace události kódem třídy
- **Harmonizace zápisu adresních atributů**
- **Geokódování**



Agregace dat ETR

- Cílem agregace je především potlačení vlivu náhodných efektů u dat s vysokým rozlišením, které by znesnadnily analýzy a interpretace, snadnější identifikace trendů a vazeb (zejména závislostí, asociací), anonymizace dat.
- Geografická agregace:
 - Podle geometrického vymezení (100x100m, 500, 1000, 5000 m)
 - Podle administrativního uspořádání územně správních jednotek
 - Podle organizačního členění PČR (stávající způsob)
 - Jádrové vyhlazení - parametry
- Agregace podle času (podle hodin, podle kalendářního data) - využívá se i nepřesně zadaných dat. Den v týdnu.
- Agregace podle klasifikace deliktu (TSK)
- Agregace podle charakteristiky delikventa a okolností deliktu (podle věku pachatele, jeho bydliště, recidivista, drogově závislý apod.)



Anonymizace dat ETŘ

- Cílem anonymizace je poskytování dat z ETŘ (či jiných evidenčních systémů) v takové formě, aby byla zajištěna ochrana osobních (případně firemních) údajů tak, aby nedošlo k ohrožení základních lidských práv dle Listiny základních práv a svobod, zejména (ale nejen) nedotknutelnost osoby, soukromí a lidské důstojnosti, a její majetková práva
- kombinace následujících postupů:
 - Agregace dat – především územní agregace
 - Selektce pro nadlimitní počet případů
 - Náhodné roztřesení
 - Náhodné přičítání a odečítání



Datové fúze různých vnitřních datových zdrojů a data mining

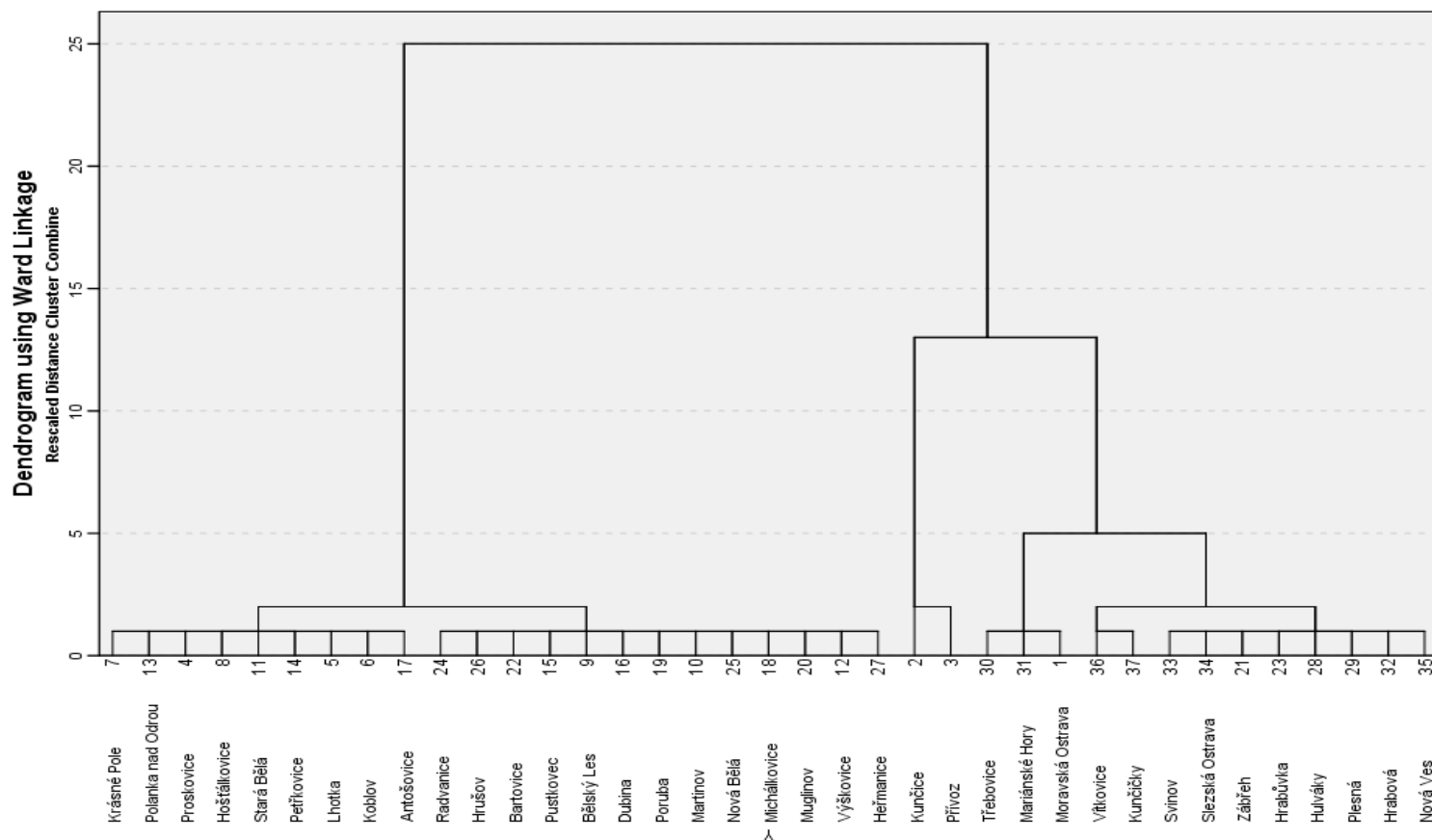
- Metody tvorby a optimalizace datového skladu,
- vytváření multidimenzionální databáze,
- Návrh OLAP, jednotlivých agregačních úrovní, dimenzí
- analytických nástrojů pro práci nad takovou databází (data mining)
- zejména shlukovací metody, metod identifikace sekvenčních vzorů, faktorové analýzy, vícenásobné lineární regrese, geografického profilování

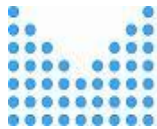


Data mining

- výběr vhodných statistických metod, vysvětlení principů použití a dokumentace ukázky použití na vybraných datech
 - Analýza hlavních komponent, ukázka pro nápad t.č. v částech města Ostrava, možnost transformace 13 indexů kriminality po kategoriích do 3 komponent zachovávajících $\frac{3}{4}$ informačního obsahu
 - Faktorová analýza explorativní, ukázka pro data SLDB2011
 - Diskriminační analýza, ukázka pro data SLDB2011
 - Shluková analýza: metoda *k*-means, ukázka pro indexy kriminality, 3 geografické shluky částí obcí podle podobnosti
 - Shluková analýza: metoda *k*-medoids,
 - Shluková analýza: Model-based clustering
 - Shluková analýza: Hierarchické aglomerativní shlukování, ukázka dendrogramu z územních jednotek s indexem kriminality
 - Shluková analýza: Hierarchické divizivní shlukování, ukázka pro data SLDB2011

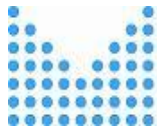
Hierarchické aglomerativní shlukování - ukázka dendrogramu z územních jednotek podle indexu kriminality





Návrh OLAP databáze pro policejní data

- Společná struktura pro PČR i MP
- Využití návrhu Metodiky harmonizace, agregace a anonymizace dat ETŘ
- Agregáční operace, 1x měsíčně
- při zpětné aktualizaci opravné (záporné a kladné) agregace
- integrace různých interních datových zdrojů – zamezit duplicitě, rozdíly v popisu



Dimenze OLAP databáze pro policejní data

- Administrativních územních jednotek (část obce, obce, ORP, okres, kraj), vč. verzování - pro místo činu, bydliště pachatele a bydliště oběti
- Kvadrantů (100x100m, 500 m, 1 km, 5 km) - pro místo činu, bydliště pachatele a bydliště oběti
- Policejní jednotky (řešící případ)
- Datum (den, typ dne, týden, měsíc, kvartál, rok)
- Čas (hodina, noc/den, ráno/večer)
- Objasněnost
- TSKU (třída → nadtřída → kategorie), rozšířen o 370 položek zahrnujících přestupky dle 30 paragrafů
- typ objektů – nový číselník
- typ podobjektů – nový číselník - 87 podobjektů
- typ vozidla
- Pohlaví pachatele a oběti
- Věk pachatele a oběti (5leté kategorie, všechny širší kategorie)
- Návyky pachatele

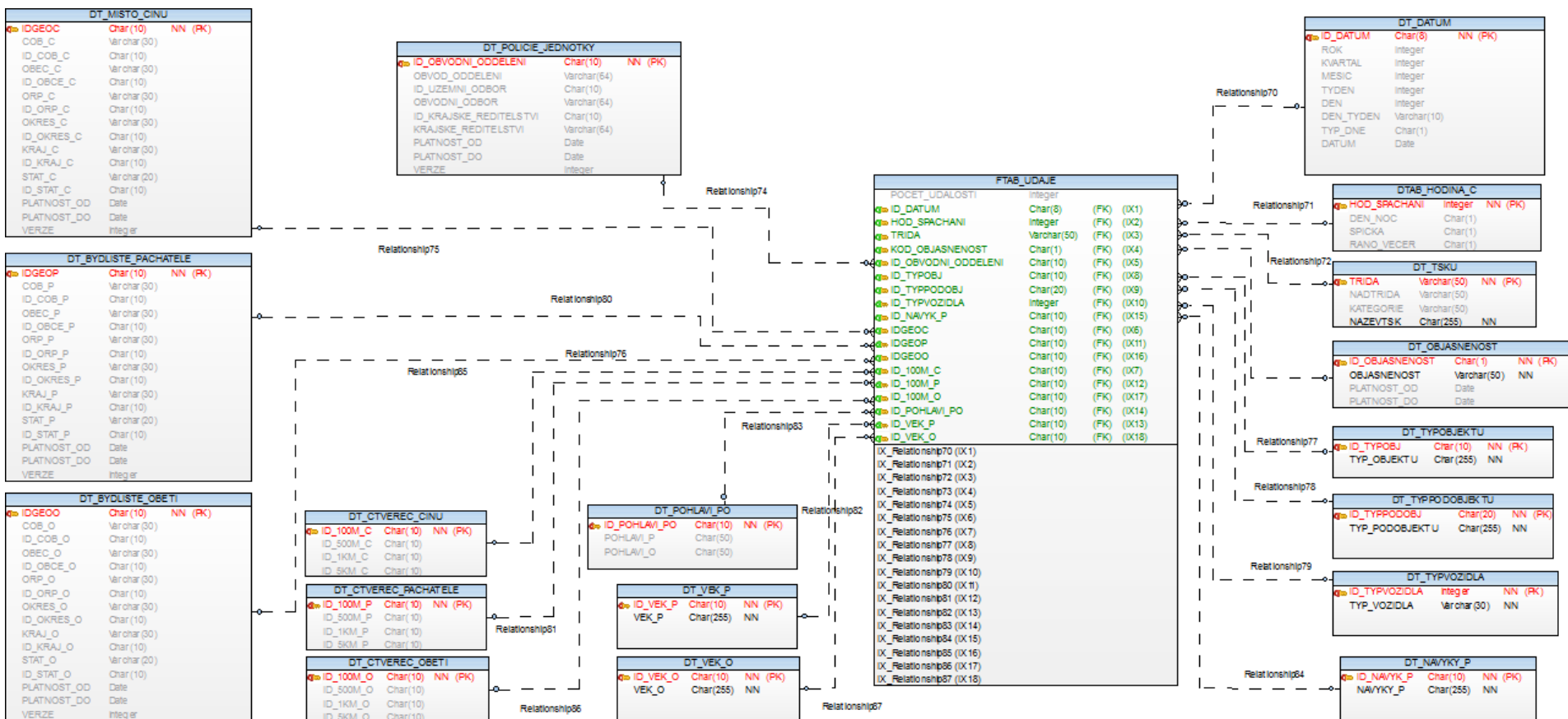


Podobjekty

OBJ2001	PODOBJEKT I - apartmá, hotelový pokoj
OBJ2002	PODOBJEKT I - ateliér
OBJ2003	PODOBJEKT I - balkón, lodžie
OBJ2004	PODOBJEKT I - buňka
OBJ2005	PODOBJEKT I - čekárna
OBJ2006	PODOBJEKT I - dílna
OBJ2007	PODOBJEKT I - garáž (řadová garáž)
OBJ2008	PODOBJEKT I - hromadné garáže (podzemní, nadzemní ...)
OBJ2009	PODOBJEKT I - herna
OBJ2010	PODOBJEKT I - chodba, schodiště
OBJ2011	PODOBJEKT I - jídelna
OBJ2012	PODOBJEKT I - kancelář
OBJ2013	PODOBJEKT I - kantýna
OBJ2014	PODOBJEKT I - klubovna
OBJ2015	PODOBJEKT I - knihovna
OBJ2016	PODOBJEKT I - kolárna, kočárkárna
OBJ2017	PODOBJEKT I - kotelna
OBJ2018	PODOBJEKT I - kuchyně
OBJ2019	PODOBJEKT I - kůlna, dřevník, stodola, stáj
OBJ2020	PODOBJEKT I - kurník, králíkárna, chlívek, včelín
OBJ2021	PODOBJEKT I - laboratoř
OBJ2022	PODOBJEKT I - maringotka
OBJ2023	PODOBJEKT I - ordinace
OBJ2024	PODOBJEKT I - pasáž, podchod (průchod, průjezd)
OBJ2025	PODOBJEKT I - počítačové pracoviště (specializované, s větší koncentrací VT)
OBJ2026	PODOBJEKT I - pokoj (ložnice, dětský, obývací, pracovna)
OBJ2027	PODOBJEKT I - požární zbrojnice
OBJ2028	PODOBJEKT I - prádelna, sušárna
OBJ2029	PODOBJEKT I - přístřešek, přístavek
OBJ2030	PODOBJEKT I - půda
OBJ2031	PODOBJEKT I - recepce
OBJ2032	PODOBJEKT I - rezidence
OBJ2033	PODOBJEKT I - sестerna, lékařský pokoj
OBJ2034	PODOBJEKT I - sklad
OBJ2035	PODOBJEKT I - sklep
OBJ2036	PODOBJEKT I - směnárna
OBJ2037	PODOBJEKT I - spíž, komora
OBJ2038	PODOBJEKT I - staveniště
OBJ2039	PODOBJEKT I - skladiště (úložiště) stavebního a podobného materiálu (oplocený areál)



Návrh OLAP databáze pro policejní data





Datová fúze zájmových vnějších zdrojů dat a data mining

- Mapování externích zdrojů dat potřebné pro lokalizaci prováděnou v analytické databázi
- vyhledávání různých zdrojů informací o zájmových objektech, jejich popis z hlediska struktury, kvality, sémantiky, režimu poskytování (dostupnost, frekvence, podmínky atd.), jejich vzájemné porovnání.
- referenční objekty, kterými rozumíme objekty, využívané policií při popisu lokalizace, jako jsou obchody a obchodní centra, restaurační zařízení, zábavní podniky, čerpací stanice, zastávky veřejné dopravy, herny, zastavárny, sběrný druhotných surovin apod. Tyto objekty také často představují místa se vztahem ke kriminalitě (zvýšená koncentrace deliktů v nich či v jejich okolí, nebo možná existence příčinné či souvislostní vazby – např. tržnice, sběrný surovin či zastavárny).
- tvorba doporučení pro výběr vhodných zdrojů a postupů ETL pro jejich systematické vytěžování a zabudování do standardních procesů aktualizace prostorových dat
- tvorba komplexní integrované geodatabáze
- příprava OLAP pro externí data, možnost analytického využití, data mining



Zdroje externích dat (mimo PČR a MP)

- Obchodní centra, obchody
- Restaurační zařízení, zábavní podniky
- Čerpací stanice
- Zastávky veřejné dopravy
- Herny
- Zastavárny,
- Sběrný druhotných surovin
- Veřejné rejstříky - spolkový rejstřík, nadační rejstřík, rejstřík ústavů, obchodní rejstřík atd.
- CEDA POI
- Ubytovny
- Bankomaty
- Školy
- Zdravotnická zařízení
- Národní registr léčby uživatelů drog
- Evidence krajské hygienické stanice a neziskových organizací
- Informační systém infekčních nemocí, Informační systém Zemřelí (IS ZEM),
- Národní registr hospitalizovaných (NRHOSP)
- SLDB
- Data Úřadu práce – uchazeči, zaměstnavatelé



OLAP databáze pro data ETŘ - město Kolín z let 2013-4

- datová kostka má 27893 agregačních buněk
- datum neurčeno u 358, s přesností pouze na rok určeno 835, 2651 buněk s přesností pouze na měsíc a zbytek je určen na úroveň dne
- Hodina neurčena u 6649 buněk
- TSK chybí u 9 buněk
- Geografické umístění je zajištěno pouze přes místo činu. Celkem 8471 buněk (zastupujících 15450 případů) nemá určenou geografickou polohu (na část obce).
- Celkem 8 stometrových buněk vykazuje více než 100 událostí
- Typy objektů - nejvíce zastávky autobusů (mimo MHD) – celkem 1926. Následují delikty na čerpací stanici (726), motorová vozidla (534+184), rodinné domky (534), rekreační objekty (264), prodejny (168), sportovní objekty (160)



OLAP databáze pro ETŘ - město Ostrava z let 2010-1

- datová kostka má 81187 agregačních buněk
- datum neurčeno pouze u 1151, s přesností pouze na rok určeno 4145, 11183 buněk s přesností pouze na měsíc a zbytek je určen na úroveň dne.
- 36049 buněk nemá ve výsledku určenou hodinu
- Typy objektů - události vztažené k motorovému vozidlu (12693). Následují supermarkety (1677 případů), restaurace a bary (1000), byty (evidovaných 575 případů), zastávky MHD (471), rodinné domky (442), na čerpací stanici (385)
- Věk pachatele není evidován u 49702 případů. Ze známých pachatelů (podezřelých) je nejčastější věkový interval 20-25 let (6059 případů), následuje 25-30 let (5285 případů) a 30-35 let (5032 případů), dále 35-40 let (3985 případů).

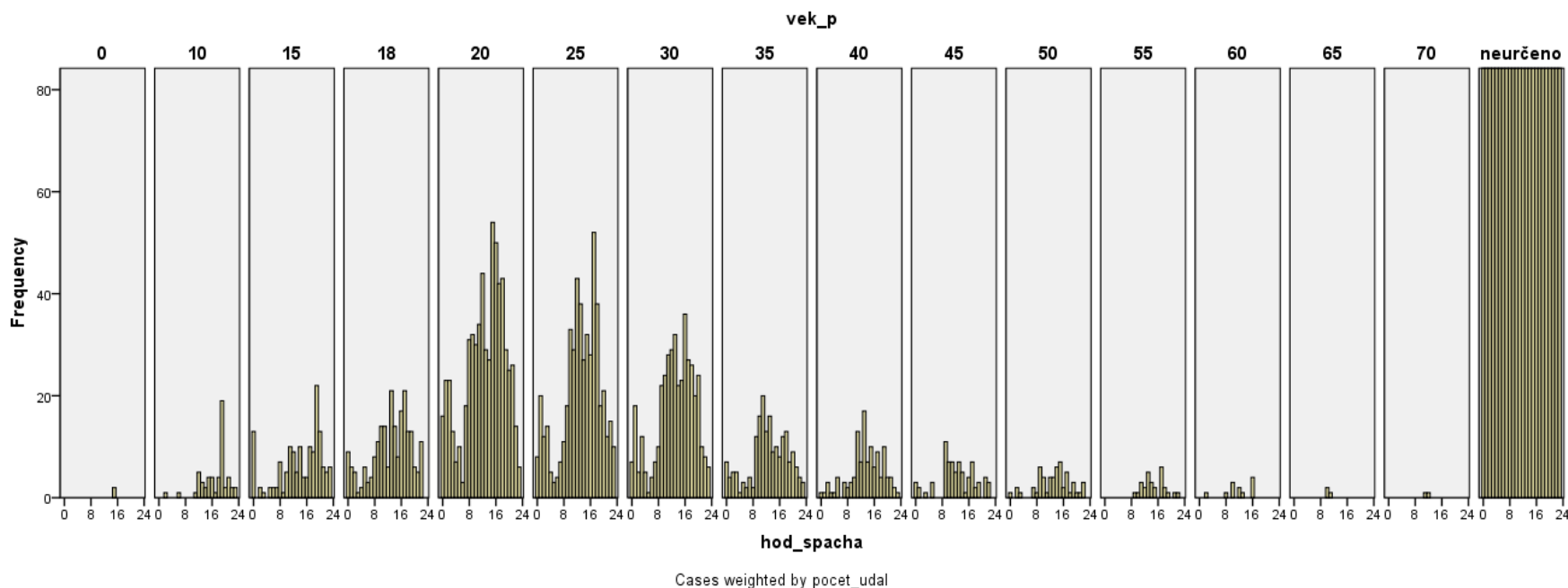
Ukázka implementace v SPSS

- vytvoření OLAP kostky
- vhodné nástroje pro manipulaci s daty (např. Pivoting trays), definování statistických ukazatelů, různých výpočtů a porovnání

OLAP Cubes							
vek_p 15-18							
typ_objekt		Sum	N	Mean	Std. Deviation	% of Total Sum	% of Total N
pocet_udal	CÍRKEVNÍ OBJEKT - hřbitov, hrobka	1	1	1,00	.	,0%	,0%
	DOPRAVNÍ PROSTREDEK - PHM	2	2	1,00	,000	,0%	,0%
	DOPRAVNÍ PROSTREDEK - součástky motorových vozidel	13	7	1,86	1,069	,1%	,1%
	DOPRAVNÍ PROSTREDEK - věci uložené v motorovém vozidle	9	9	1,00	,000	,1%	,1%
	MIMO OBJEKT - zastávka MHD	4	4	1,00	,000	,0%	,0%
	neurčeno	185	123	1,50	,670	1,7%	1,2%
	PRODEJNA - supermarket (hypermarket, obchodní dům)	2	2	1,00	,000	,0%	,0%
	Total	216	148	1,46	,674	1,9%	1,4%

Ukázka implementace v SPSS

- Distribuce krádeží v jednotlivých hodinách dne podle věku pachatele (2010-1)



Možnosti využití

- Agregace ve dnech, hodinách

Časový průběh znečišťování veřejného prostranství v Ostravě v průběhu týdne

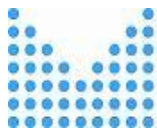
2012

	0 h	1 h	2 h	3 h	4 h	5 h	6 h	7 h	8 h	9 h	10 h	11 h	12 h	13 h	14 h	15 h	16 h	17 h	18 h	19 h	20 h	21 h	22 h	23 h	0 h
Pondělí	9	12	10	6	4	3	6	14	56	98	82	58	75	93	90	42	40	30	11	33	36	68	33	24	
Úterý	7	9	9	3	1	6	4	14	47	86	80	56	61	93	49	43	38	16	14	32	52	87	34	15	
Středa	16	13	11	5	5	5	2	18	66	91	72	49	77	101	83	60	25	30	15	30	40	47	32	13	
Čtvrtek	10	17	5	4	3	6	3	22	57	82	75	79	67	89	62	44	31	18	6	24	42	67	47	24	
Pátek	24	15	15	5	7	5	1	11	50	82	98	57	51	94	91	44	25	19	5	23	64	64	51	33	
Sobota	25	26	14	16	7	4	8	12	56	86	72	50	66	66	62	52	17	18	15	30	43	60	37	29	
Neděle	27	35	19	19	12	4	6	16	44	75	94	69	75	73	70	32	23	15	12	23	58	55	37	26	

2013

	0 h	1 h	2 h	3 h	4 h	5 h	6 h	7 h	8 h	9 h	10 h	11 h	12 h	13 h	14 h	15 h	16 h	17 h	18 h	19 h	20 h	21 h	22 h	23 h	0 h
Pondělí	22	22	15	9	1	1	4	19	74	90	82	65	83	98	63	42	24	30	11	40	58	53	32	13	
Úterý	24	19	14	11	4	8	4	22	74	74	81	58	67	83	81	55	30	28	19	51	73	70	27	23	
Středa	23	19	8	5	2	6	6	20	97	106	93	59	94	102	79	45	25	39	12	33	81	55	30	10	
Čtvrtek	10	19	12	9	2	3	7	32	56	100	82	48	70	75	70	48	29	31	6	56	70	79	39	35	
Pátek	15	13	8	11	5	3	4	25	82	86	105	59	81	107	75	44	44	23	11	29	64	68	30	25	
Sobota	16	27	25	12	13	7	8	36	80	97	99	68	58	76	68	50	32	17	14	31	68	81	41	25	
Neděle	37	33	21	15	9	7	10	32	83	86	107	54	63	81	69	58	36	25	13	36	63	77	33	18	

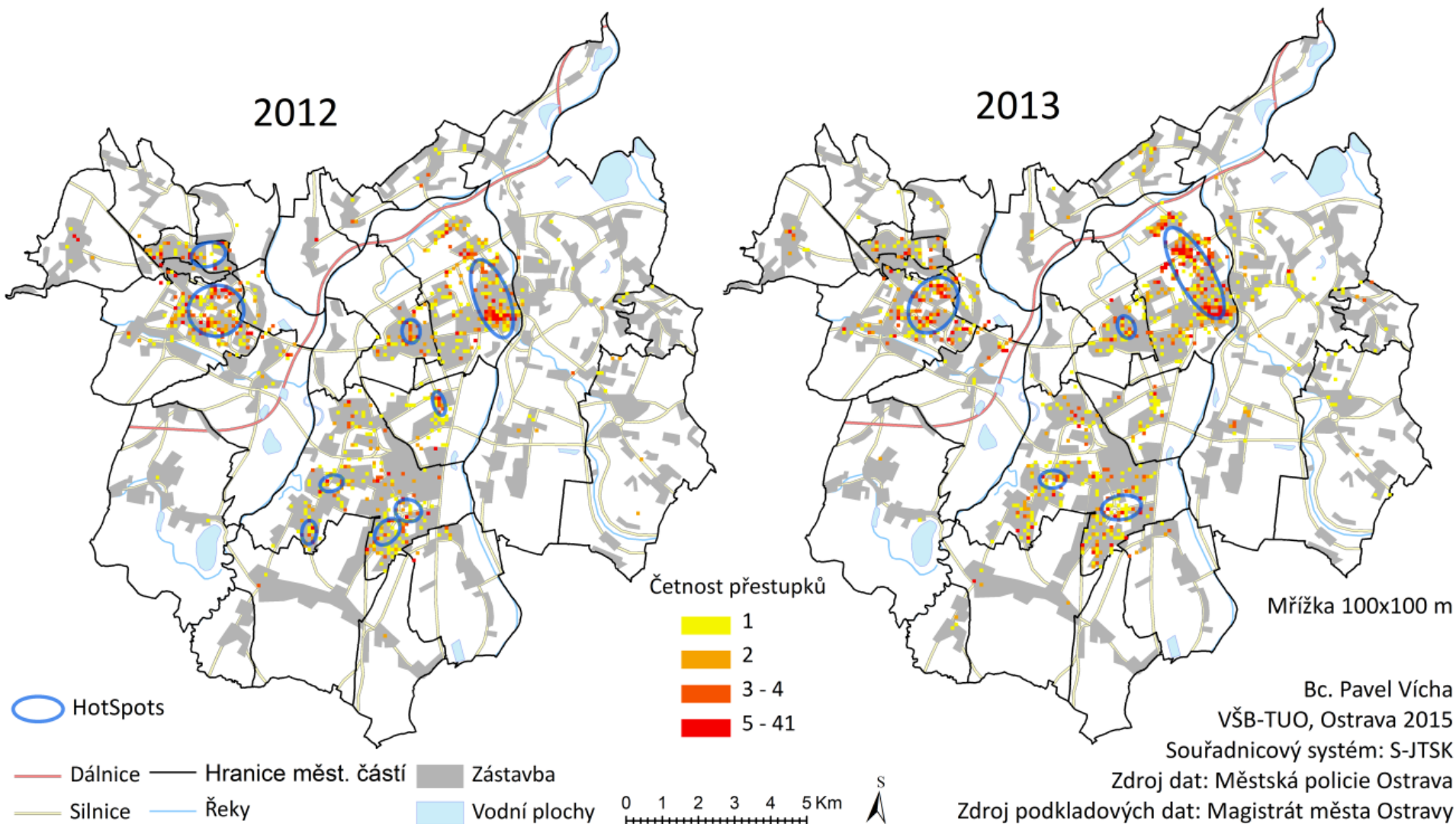
- Prostorová distribuce, změny v čase, trendy, indikace problémových míst
- Kombinace více faktorů



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



ČETNOST PŘESTUPKŮ RUŠENÍ NOČNÍHO KLIDU V OSTRAVĚ V LETECH 2012 A 2013





Analytické využití OLAP – vztahy mezi demografií, nezaměstnaností a kriminalitou (Ostrava, 2010)

- Posuzování intenzity – předběžné výsledky:
 - vyšší index kriminality – místa s vysokým podílem dětí, nezaměstnaných i dlouhodobě nezaměstnaných
 - Krádeže vloupání – místa s vysokým podílem dětí, nezaměstnaných i dlouhodobě nezaměstnaných
 - násilné t.č. – nejsou tam, kde je hodně seniorů.
 - krádeže – místa s vysokým podílem seniorů a dlouhodobě nezaměstnaných
 - přestupky par.50 – místa s vyšším výskytem dětí i seniorů a dlouhodobě nezaměstnaných.





Correlations

Pearson Correlation

	pod_001 4	pod_650 0	pod_uc_op v	pod_ucAB C	pod_uce1 2	in_kri mi	in_kr_tsk a	in_kr_tsk d	in_kr_tsk e	in_kr_tspr0 4
pod_0014	1	-,462**	,488**	,423**	,311**	,170*	,036	,495**	,062	,162*
pod_6500	-,462**	1	,004	-,245**	-,239**	,096	-,259*	,062	,254**	,167*
pod_uc_op v	,488**	,004	1	,232**	,234**	,176*	,045	,341**	,084	,158
pod_ucAB C	,423**	-,245**	,232**	1	,570**	,130	,072	,164	,138	,037
pod_uce12	,311**	-,239**	,234**	,570**	1	,351**	,184	,464**	,250**	,351**
in_krimi	,170*	,096	,176*	,130	,351**	1	,906**	,662**	,919**	,752**
in_kr_tska	,036	-,259*	,045	,072	,184	,906**	1	,791**	,892**	,711**
in_kr_tskd	,495**	,062	,341**	,164	,464**	,662**	,791**	1	,389**	,561**
in_kr_tske	,062	,254**	,084	,138	,250**	,919**	,892**	,389**	1	,701**
in_kr_tspr0 4	,162*	,167*	,158	,037	,351**	,752**	,711**	,561**	,701**	1

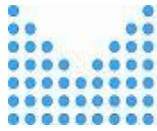
** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).



Doporučení k integraci dat a opatření ke zlepšení situace v evidenci prostorových dat

- Změny a úpravy informačních systémů
 - Sjednocení evidence adres napříč IS,
 - Zlepšení identifikace územních jednotek (u MP) - navázání na oficiální územní číselníky
 - evidence souřadnic přímo měřených v terénu GNSS (bodové i trasy)
 - časová platnost u číselníků - používat aktuální údaje a současně při práci s historickými údaji aplikovat časově konzistentní číselníky
 - Zobecnění a sjednocení evidence organizací a jejich provozoven napříč informačními systémy
 - Zobecnění a sjednocení evidence objektů. Hierarchický systém evidence objektů, evidence typu objektu, jeho názvu, případných podtypů, vymezením časové platnosti, souřadnic vztažného bodu (případně polygon), adresy (stavební objekty)
 - Klasifikaci přestupků u MP zobecnit, ne paragrafové znění zákona
 - vhodné datové typy a vhodné identifikátory u MP
- Zavedení doplňkových systémů lokalizace v území (mimo použití přijímačů GNSS)
- Opatření při pořizování dat
- Opatření při zpracování dat – geokódování
- Opatření při analýze dat a tvorbě výstupů - využití údaje o přesnosti lokalizace



MINISTERSTVO VNITRA
ČESKÉ REPUBLIKY



Děkuji za pozornost.

Dotazy?