

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra matematiky

Bakalářská práce

Aplikace vybraných demografických metod

Plzeň, 2009

Pavel Hájek

Zadání bakalářské práce – vložený list.

Prohlášení

Předkládám tímto k posouzení a obhajobě bakalářskou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně a výhradně s použitím zdrojů, jejichž úplný seznam je její součástí.

V Plzni dne 29.5.2009

.....

Poděkování

Chtěl bych na tomto místě poděkovat všem osobám, podílejícím se na vzniku této bakalářské práce. Především vedoucímu bakalářské práce Ing. Karlu Janečkovi za odbornou pomoc a vedení a v neposlední řadě také svým rodičům za soustavnou podporu.

Abstrakt

Předložená práce je dělena do tří hlavních tématických částí. První část této práce se zabývá návrhem vhodné databázové struktury pro uložení dat obsahující informace o poloze a značení sídelních jednotek a následným vytvořením této databáze. Jedná se o zhotovení základních, daty nenaplněných, tabulek SIDLA a SIDLA_MULLER, jejichž návrhem, vytvořením a popisem se zabývá první kapitola. Druhá část se věnuje již samotnému naplnění a aktualizaci vytvořených tabulek. V rámci druhé kapitoly byla konkrétně tabulka SIDLA aktualizována daty poskytovanými Českým statistickým úřadem, tabulka SIDLA_MULLER pak daty z prací zabývajících se lokalizací Müllerových rukopisných map Čech. Závěrečná část, kapitola třetí, obsahuje vybrané a provedené demografické statistiky a analýzy vycházející z dat uložených ve výše zmíněných tabulkách. Přitom vznikla řada dalších potřebných tabulek k realizaci těchto analýz a statistik, jejichž úplný přehled a popis je součástí předložené práce.

Klíčová slova

databáze, tabulka, data, PostgreSQL, Müllerovy rukopisné mapy krajů, analýza, statistika

Abstract

Shown work is divided into three main theme parts. The first part of this work concerns with a proposal for suitable database structure to store data that contain information about placing and marking of settlement units and with following creation of this database. It is going about basic, data free tables SIDLA and SIDLA_MULLER of which proposal, creation and description is concerned in first chapter. The second part engages in filling and actualization of created tables already. In the frame of the second chapter was the table SIDLA actualized with the data provided by Czech Statistic Office and the table SIDLA_MULLER with data acquired from the works focused on the location of Müller's holograph maps of Czech. The closing part, the third chapter, contains chosen and proceeded demographic statistics and analysis coming out from the data stored in higher mentioned tables. Along with this resulted from it a bunch of other tables necessary for implementation of this analysis and statistics of which complete overview and description is a part of enclosed work.

Keywords

database, table, data, PostgreSQL, Müller's holographic maps of regions, analyse, statistic

Obsah

| | |
|--|----|
| Seznam obrázků..... | 6 |
| Seznam tabulek..... | 8 |
| Seznam příkazů a vzorců..... | 9 |
| Zkratky a značky..... | 10 |
| Typografická konvence..... | 10 |
| Úvod..... | 11 |
| 1. Vytvoření databáze sídel..... | 12 |
| 1.1 Struktura databáze..... | 12 |
| 1.2 Volba databázového systému..... | 14 |
| 2. Naplnění a aktualizace databáze..... | 15 |
| 2.1 Prvotní plnění databáze..... | 15 |
| 2.2 Aktualizace databáze zpřesněnými souřadnicemi sídel..... | 19 |
| 2.3 Uložení dat z Müllerových map..... | 24 |
| 3. Aplikace metod..... | 31 |
| 3.1 Polohově závislé..... | 31 |
| 3.1.1 Rozdíly souřadnic mezi entitami v tabulce SIDLA a SIDLA_MULLER..... | 31 |
| 3.1.2 Rozdíly souřadnic mezi sídly, která se v tabulce SIDLA_MULLER vyskytují více než jednou..... | 39 |
| 3.2 Polohově nezávislé, atributové..... | 45 |
| 3.2.1 Úplná shoda názvů sídel..... | 45 |
| 3.2.2 Rozdíl německých názvů sídel v rámci překrývajících se mapových listů..... | 47 |
| 3.2.3 Poměr typů sídel na Müllerových mapách..... | 49 |
| 3.2.4 Vývoj typů sídel mezi počátkem 18. století a současností..... | 61 |
| Závěr..... | 69 |
| Seznam příloh..... | 70 |
| Použitá literatura..... | 78 |

Seznam obrázků

| | |
|---|----|
| Obrázek 1: Schéma struktury databáze..... | 12 |
| Obrázek 2: Ukázka zdrojového souboru sidla.stx..... | 16 |
| Obrázek 3: Náhled souboru sidla-insert.sql vytvořený pro naplnění databáze prvotními daty | 17 |
| Obrázek 4: Ukázka DB sídel naplněné prvotními daty..... | 18 |
| Obrázek 5: Část aktualizací souboru sidla aktual.stx..... | 19 |
| Obrázek 6: Část aktualizací souboru sidla smaz.stx..... | 20 |
| Obrázek 7: Část souboru pro aktualizaci DB sídel vygenerovaný programem ze souborů sidlaXaktual.stx..... | 21 |
| Obrázek 8: Výpis z tabulky SIDLA po aktualizaci ze souborů sidlaXaktual.stx..... | 22 |
| Obrázek 9: Vyhledání zbylých sídel s atributem zmeneno = 2 v kompletně aktualizované tabulce SIDLA..... | 23 |
| Obrázek 10: Ukázka souboru k naplnění tabulky SIDLA_MULLER..... | 25 |
| Obrázek 11: Ukázka souboru pro vkládání záznamů do tabulky SIDLA_MULLER..... | 26 |
| Obrázek 12: Tabulka SIDLA obsahující osmimístná id..... | 27 |
| Obrázek 13: Vyobrazení tabulky SIDLA_MULLER po aktualizaci bez udání typu sídel..... | 28 |
| Obrázek 14: Soubor se sídly rozdělenými dle typu (zde vsi se zámky)..... | 29 |
| Obrázek 15: Ukázka dat, díky kterým se vloží do tabulky SIDLA_MULLER informace o typu sídla..... | 29 |
| Obrázek 16: Ukázka plně naplněné tabulky SIDLA_MULLER..... | 30 |
| Obrázek 17: Ukázka výstupu z tabulky ROZDIL_SIDLA_MULLER..... | 32 |
| Obrázek 18: Část tabulky ROZDIL_SIDLA_MULLER_UNIKATNI..... | 33 |
| Obrázek 19: Ukázka tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 35 |
| Obrázek 20: Částečná ukázka tabulky ODCHYLKY_SIDLA_MULLER..... | 38 |
| Obrázek 21: Ukázka tabulky POCET_SHODNYCH_SIDELO..... | 40 |
| Obrázek 22: Ukázka tabulky ROZDIL_V_MULLEROVI..... | 41 |
| Obrázek 23: Vyobrazení části tabulky ROZDIL_V_MULLEROVI_UNIKATNE..... | 42 |
| Obrázek 24: Část tabulky SHODNE_NAZVY..... | 46 |
| Obrázek 25: Část přehledu tabulky VYBER_JINYCH_NEMECKYCH..... | 47 |
| Obrázek 26: Ukázka výběru zobrazující rozdílnost typu sídel u duplicit..... | 49 |
| Obrázek 27: Část tabulky MULLER_UNIQUE_ZAZNAMY..... | 50 |
| Obrázek 28: Ukázka tabulky POMER_UNIQUE..... | 51 |
| Obrázek 29: Část vyobrazení tabulky MULLER_REDUNDANTNI_ZAZNAMY..... | 53 |
| Obrázek 30: Zobrazení tabulky POMER_REDUNDANTNI..... | 54 |
| Obrázek 31: Ukázka tabulky MULLER_UNIQUE_I_REDUNDANTNI..... | 55 |
| Obrázek 32: Přehled tabulky POMER_VSECH..... | 56 |
| Obrázek 33: Ukázka tabulky MULLER_ZANIKLA_SIDLA..... | 58 |
| Obrázek 34: Přehled tabulky POMER_ZANIKLA..... | 59 |
| Obrázek 35: Finální podoba tabulky SIDLA pro tuto práci..... | 63 |
| Obrázek 36: Přehled tabulky TYP_CSU..... | 64 |
| Obrázek 37: Finální struktura výchozích tabulek pro analýzy..... | 64 |
| Obrázek 38: Přehled tabulky SROVNANI_OBCE..... | 66 |
| Obrázek 39: Přehled tabulky SROVNANI_KU..... | 66 |
| Obrázek 40: Přehled tabulky SROVNANI_MCAST..... | 66 |
| Obrázek 41: Přehled tabulky SROVNANI_COBE..... | 67 |
| Obrázek 42: Přehled tabulky SROVNANI_COB..... | 67 |
| Obrázek 43: Přehled tabulky SROVNANI_ZSJ..... | 67 |
| Příloha A: Přehled krajů Čech znázorněných na Müllerových rukopisných mapách [9]..... | 70 |
| Příloha 1: Výběr odchylek tabulky ROZDIL_SIDLA_MULLER_UNIKATNI..... | 71 |

| | |
|--|----|
| Příloha 2 : Výběr odchylek tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 72 |
| Příloha 3 : Výběr odchylek z tabulky ROZDIL_V_MULLEROVI_UNIKATNE..... | 73 |

Seznam tabulek

| | |
|--|----|
| Tabulka 1: Popis databázové tabulky SIDLA..... | 12 |
| Tabulka 2: Popis databázové tabulky SIDLA_MULLER..... | 13 |
| Tabulka 3: Popis databázové tabulky KRAJ..... | 13 |
| Tabulka 4: Popis databázové tabulky TYP_SIDEK..... | 14 |
| Tabulka 5: Popis databázové tabulky NALEZITOST..... | 14 |
| Tabulka 6: Přehled souborů definičních bodů pro užití v DB..... | 15 |
| Tabulka 7: Popis tabulky ROZDIL_SIDLA_MULLER..... | 31 |
| Tabulka 8: Popis tabulky ROZDIL_SIDLA_MULLER_UNIKATNI..... | 33 |
| Tabulka 9: Popis tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 34 |
| Tabulka 10: Přehled odchylek tabulky ROZDIL_SIDLA_MULLER_UNIKATNI..... | 36 |
| Tabulka 11: Přehled odchylek tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 37 |
| Tabulka 12: Popis tabulky POCET_SHODNYCH_SIDEK..... | 39 |
| Tabulka 13: Popis tabulky ROZDIL_V_MULLEROVI..... | 41 |
| Tabulka 14: Popis tabulky ROZDIL_V_MULLEROVI_UNIKATNE..... | 42 |
| Tabulka 15: Přehled odchylek tabulky ROZDIL_V_MULLEROVI_UNIKATNE..... | 43 |
| Tabulka 16: Popis tabulky SHODNE_NAZVY..... | 45 |
| Tabulka 17: Popis tabulky VYBER_JINYCH_NEMECKYCH..... | 47 |
| Tabulka 18: Popis tabulky MULLER_UNIQUE_ZAZNAMY..... | 50 |
| Tabulka 19: Popis tabulky POMER_UNIQUE..... | 51 |
| Tabulka 20: Poměr neredundantních sídel v tabulce POMER_UNIQUE..... | 52 |
| Tabulka 21: Popis tabulky MULLER_REDUNDANTNI_ZAZNAMY..... | 53 |
| Tabulka 22: Popis tabulky POMER_REDUNDANTNI..... | 54 |
| Tabulka 23: Poměr všech sídel v tabulce POMER_REDUNDANTNI..... | 54 |
| Tabulka 24: Popisek tabulky POMER_VSECH..... | 56 |
| Tabulka 25: Poměr všech sídel v tabulce POMER_VSECH..... | 56 |
| Tabulka 26: Popis tabulky MULLER_ZANIKLA_SIDLA..... | 57 |
| Tabulka 27: Popis tabulky POMER_ZANIKLA..... | 58 |
| Tabulka 28: Poměr již zaniklých sídel v tabulce SIDLA_MULLER..... | 59 |
| Tabulka 29: Celkový přehled zastoupení typů sídel..... | 59 |
| Tabulka 30: Přehled užitých databázových souborů..... | 61 |
| Tabulka 31: Popis tabulky TYP_CSU..... | 64 |
| Tabulka 32: Konečná podoba tabulky SIDLA..... | 65 |
| Tabulka 33: Součet sídel dle kategorií ČSÚ..... | 67 |

Seznam příkazů a vzorců

| | |
|--|----|
| Příkaz 1: Kontrola vložených dat do tabulky SIDLA..... | 22 |
| Příkaz 2: Sekvence pro generování hodnoty id pro tabulku SIDLA_MULLER..... | 26 |
| Příkaz 3: Sql dotaz pro vytvoření tabulky ROZDIL_SIDLA_MULLER..... | 31 |
| Příkaz 4: Výběr unikátních sídel z tabulky ROZDIL_SIDLA_MULLER..... | 32 |
| Vzorec 1: Výpočet polohové odchylky Δ_{xy} pro každé sídlo uložené v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI..... | 33 |
| Vzorec 2: Výpočet průměrné polohové odchylky Δ pro záznamy v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI..... | 33 |
| Příkaz 5: Výběr duplicitních sídel z tabulky ROZDIL_SIDLA_MULLER..... | 34 |
| Vzorec 3: Výpočet polohové odchylky Δ_{xy} pro každé sídlo uložené v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 35 |
| Vzorec 4: Výpočet průměrné polohové odchylky Δ pro záznamy v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI..... | 35 |
| Příkaz 6: Výběr hodnot odchylka z tabulky ROZDIL_SIDLA_MULLER_UNIKATNI dle intervalů..... | 35 |
| Příkaz 7: Výběr hodnot odchylka z tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI dle intervalů..... | 36 |
| Příkaz 8: Tvorba tabulky ODCHYLKY_SIDLA_MULLER..... | 37 |
| Příkaz 9: Vytvoření tabulky POCET_SHODNYCH_SIDEL..... | 39 |
| Příkaz 10: Tvorba tabulky ROZDIL_V_MULLEROVI..... | 40 |
| Příkaz 11: Vytvoření tabulky ROZDIL_V_MULLEROVI_UNIKATNE..... | 42 |
| Vzorec 5: Výpočet střední polohové odchylky pro data uvedená v tabulce ROZDIL_V_MULLEROVI_UNIKATNE..... | 43 |
| Příkaz 12: Vytvoření tabulky SHODNE_NAZVY..... | 45 |
| Příkaz 13: Tvorba tabulky VYBER_JINYCH_NEMECKYCH..... | 47 |
| Příkaz 14: Výběr rozdílných typů sídel v rámci duplicitních sídel v tabulce SIDLA_MULLER | 49 |
| Příkaz 15: Tvorba tabulky MULLER_UNIQUE_ZAZNAMY..... | 50 |
| Příkaz 16: Tvorba tabulky POMER_UNIQUE..... | 51 |
| Příkaz 17: Tvorba tabulky MULLER_REDUNDANTNI_ZAZNAMY..... | 52 |
| Příkaz 18: Tvorba tabulky POMER_REDUNDANTNI..... | 53 |
| Příkaz 19: Vytvoření tabulky MULLER_UNIQUE_I_REDUNDANTNI..... | 55 |
| Příkaz 20: Tvorba tabulky POMER_VSECH..... | 55 |
| Příkaz 21: Vytvoření tabulky MULLER_ZANIKLA_SIDLA..... | 57 |
| Příkaz 22: Tvorba tabulky POMER_ZANIKLA..... | 58 |
| Příkaz 23: Finální aktualizace tabulky SIDLA..... | 62 |
| Příkaz 24: Vzor pro tvorbu tabulky SROVNANI_OBCE a jí podobných..... | 65 |
| Příkaz 25: Vytvoření výběrů z tabulek SROVNANI_XXX..... | 66 |

Zkratky a značky

| | |
|---------|--|
| ČR | Česká republika |
| ČSÚ | Český statistický úřad |
| DB | databáze |
| ÚIR-ZSJ | Územně identifikační registr základních sídelních jednotek |

Typografická konvence

- Data Definition Language příkazy SQL - Courier New, vel. 9

`CREATE table; INSERT INTO sidla`

- Názvy sloupců databázové tabulky - Arial, vel. 12

nazev, zmeneno

- Názvy jednotlivých tabulek v databázi - Times New Roman, vel. 12

SIDLA, SIDLA_MULLER

Úvod

Dle [1] je demografie vědní obor, který zkoumá reprodukci lidských populací. V mojí práci se zaměřím spíše na jiný, než klasický způsob vnímání tohoto oboru, a to na oblast umístění a vývoje lokalizace sídel na území České republiky, respektive na části území Čech.

K tomuto úkolu jsou však nutná podkladová data. Pro svojí práci jsem si zvolil data získaná z lokalizace rukopisných Müllerových map Čech zhotovených v období 1712 - 1718 (více [2], [3] a [9]). K porovnání s nynějším stavem byla převzata data z ČSÚ, aktuální k datu 2.1.2008 [4].

Tyto informace bylo nutno separovat a uchovat ve vhodné databázové struktuře, čímž se zabývá první část mé bakalářské práce. Cílem bylo navrhnout vhodnou strukturu i pro navazující projekty, nejen pro účel této práce.

Druhá kapitola se vztahuje k naplnění databáze daty, to znamená tvorba aktualizčních programů v objektově orientovaném programovacím jazyce Java a následné vložení dat do struktury databáze.

Ve třetí části přistoupím již k vybraným analýzám a statistickým postupům použitelných pro podkladová data. Tyto metody se musely zvolit nejen s ohledem na strukturu a náplň DB, ale i na využitelnost a relevantnost výstupu z analýz. Těchto analýz bylo provedeno šest, dvě polohově závislé, čtyři atributové.

Závěrem se věnuji zhodnocení průběhu realizace databáze, jejího plnění daty a výsledků z analýz. K tomu patří i určení vypovídajících hodnot, relevantnosti těchto výsledků a možnosti jejich použití či k pozdějšímu rozšíření.

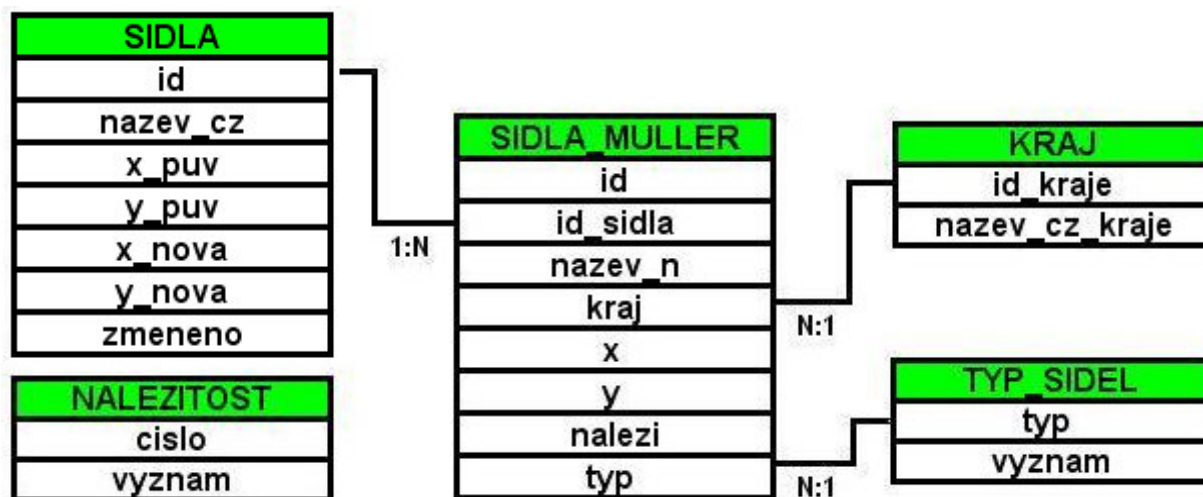
1. Vytvoření databáze sídel

Základními požadavky pro tuto práci bylo vytvoření datové struktury, její naplnění daty a provedení vybraných analýz z tohoto souboru dat. Výsledkem těchto požadavků je vytvoření normalizovaného datového modelu (viz. Obrázek 1).

1.1 Struktura databáze

Na použitý datový model byly kladeny nároky rozšiřitelnosti, efektivity a konzistence. Vzhledem k těmto nárokům byla vytvořena struktura znázorněná na Obrázku 1.

Obrázek 1: Schéma struktury databáze



Veškeré popisy atributů v tabulkách jsou popsány v Tabulkách 1 až 5.

Tabulka 1: Popis databázové tabulky SIDLA

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| id | bigint | jinak také označen jako primární klíč, sloužící pro přístup k datům (entitám tabulky) |
| nazev_cz | text | název sídla v českém jazyce |
| x_puv | double precision | původní X-ová souřadnice sídla poskytnutá ČSÚ |
| y_puv | double precision | původní Y-ová souřadnice sídla poskytnutá ČSÚ |
| x_nova | double precision | aktualizovaná X-ová souřadnice určená na ZM10 [2], [3] |
| y_nova | double precision | aktualizovaná Y-ová souřadnice určená na ZM10 [2], [3] |
| zmeneno | integer | popis v Tabulce 5 |

Tabulka SIDLA slouží jako základní stavební část DB. Je naplněna daty získanými od ČSÚ a daty z prací studentů zabývajícími se lokalizací a zpřesňováním polohy sídel a sídelních jednotek (více o tom přímo v pracích, které se touto problematikou zabývaly, viz. [2] a [3]). Obsahuje sloupec id (primární klíč), přes který se k ní odkazují příbuzné tabulky.

Tabulka 2: Popis databázové tabulky SIDLA_MULLER

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| id | integer | označen jako primární klíč, slouží pro přístup k datům (entitám tabulky) |
| id_sidla | integer | cizí klíč pro přístup k tabulce SIDLA, ke sloupci id |
| nazev_n | text | název sídla v německém jazyce |
| kraj | text | název kraje, kterému sídlo přísluší |
| x | double precision | X-ová souřadnice po transformaci z Müllerových map, poskytnutá v rámci prací studentů [2], [3] |
| y | double precision | Y-ová souřadnice po transformaci z Müllerových map, poskytnutá v rámci prací studentů [2], [3] |
| nalezi | integer | zda sídlo leží vně (hodnota 1) či uvnitř (hodnota 0) kraje na příslušném mapovém listě |
| typ | integer | typ města (více Tabulka 4) |

Tabulka 2 slouží pro uložení dat získaných v rámci prací na lokalizaci sídel na Müllerových rukopisných mapách ([2], [3]).

Tabulka 3: Popis databázové tabulky KRAJ

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| id_kraje | text | významová zkratka názvu kraje na Müllerových mapách, cizí klíč pro přístup k tabulce SIDLA_MULLER, k sloupci kraj |
| nazev_cz_kraje | text | úplný název kraje |

Tabulka 3 slouží jako referenční tabulka pro kontrolu vstupních dat do sloupce kraj v tabulce SIDLA_MULLER, zamezuje vložení jiné, než povolené hodnoty obsažené v této tabulce.

Tabulka 4: Popis databázové tabulky TYP_SIDEL

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | kód typu sídla zobrazeného na Müllerových mapách, cizí klíč k tabulce SIDLA_MULLER, k sloupci typ |
| vyznam | text | klíč k určení typu sídla |

Tabulka 4 je opět referenční, určená pro kontrolu vstupních dat, tentokrát pro data uložená ve sloupci typ tabulky SIDLA_MULLER.

Tabulka 5: Popis databázové tabulky NALEZITOST

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|-----------------------------------|
| cislo | integer | kódové označení průběhu tvorby DB |
| vyznam | text | popis fáze tvorby DB |

Tabulka 5 je významová tabulka popisující hodnoty ve sloupci zmeneno tabulky SIDLA.

1.2 Volba databázového systému

Pro vytvoření DB byl zvolen databázový systém PostgreSQL. Přínosem databázového systému PostgreSQL je jeho BSD licence (běžná licence otevřeného zdroje, není zde žádné omezení ohledně užití kódu zdroje, více [5]) a mimo to jsou s tímto systémem na fakultě velké zkušenosti. Tato DB bude výhledově sloužit nejen pro úlohy řešené v této práci, ale bude také základem pro další navazující projekty.

Databáze je uložena na fakultním serveru s následujícími specifikacemi:

HW:

CPU Xeon 5130 2GHz

RAM 4GB

HDD 500GB

SW

užívá technologii XEN, tzn. virtuální stroj

OS Debian Linux amd64

web server Apache

Práce s databází se realizuje vzdáleným přístupem přes program pgAdmin III [6] sloužící jako rozhraní ke správě, údržbě a aktualizaci této DB. Zahrnuta byla i ochrana uložených dat (díky umístění na fakultním serveru, který využívá periodické zálohování), čímž je omezena možnost ztráty informací. Díky těmto opatřením mají přístup k databázi pouze autorizované osoby a tudíž nehrozí ztráta či znehodnocení uložených dat vnějším zásahem.

2. Naplnění a aktualizace databáze

2.1 Prvotní plnění databáze

Jako vstupní data jsou použity soubory ve formátu .DBF poskytnuté od ČSÚ [4]. Jejich přehled je znázorněn v Tabulce 6.

Tyto soubory vychází (dle [2] a [4]) z Územně identifikačního registru základních sídelních jednotek (dále ÚIR-ZSJ), což je soustava databázových číselníků jednotek územně správního, technického a sídelního členění státu do úrovně podrobnosti částí obcí, k.ú. a základních sídelních jednotek. Obsahuje jejich názvy, kódy, vzájemné vazby a doplňující informativní údaje. ÚIR-ZSJ byl vytvářen v letech 1992 – 2004 pod pověřením Ministerstva pro místní rozvoj, od března 2004 patří pod záštitu ČSÚ. Tato data jsou aktuální k datu 2.1.2008 (více o struktuře, významu a využití souborů definičních bodů - Kapitola 3.2.4).

Tabulka 6: Přehled souborů definičních bodů pro užití v DB

| Soubor | Konkrétní název souboru |
|---|-------------------------|
| definičních bodů obcí a vojenských újezdů | OBCE.DBF |
| katastrálních území | KU.DBF |
| městských obvodů a městských částí | MCAST.DBF |
| částí obcí | COB.DBF |
| částí obcí evidenčních | COBE.DBF |
| základních sídelních jednotek | ZSJ.DBF |

Výhodou těchto zdrojových souborů je, že pokrývají celé území České republiky a jsou poskytovány zdarma online na internetových stránkách ČSÚ.

Výsledkem zpracování (dle [2]) těchto souborů byl textový soubor sidla.stx, který měl následující strukturu: na každém řádku (kromě hlavičky) bylo uvedeno generované id, souřadnice Y, souřadnice X (oboje souřadnice v soustavě Jednotné trigonometrické síť katastrální) a název sídla (u kterého byla mezera v názvu reprezentovaná podtržítkem) (více viz. Obrázek 2).

Obrázek 2: Ukázka zdrojového souboru sidla.stx

```

sidla YX M O O
1      614503      1063020 Rájec
2      619751      1060460 Čičová
3      620463      1059198 Velká_Černná_nad_orlicí
4      613703      1048871 Domašín_u_Černíkovíc
5      627791      1036572 Skršice
6      603618      1037438 Jedlová_v_orlických_horách
7      609073      1041098 Hlinné_u_Dobrého
8      610321      1038393 Kamenice_u_Dobrého
9      607984      1037832 Rovné_u_Dobrého
10     617351      1032747 Běstvíny
11     613919      1036437 Domašín_u_Dobrušky
12     616225      1034778 Křovice
13     616723      1037702 Mělčany_u_Dobrušky
14     618553      1035211 Pulice
15     611114      1037894 Spáleniště
16     612012      1057912 Vyhnánov
17     617015      1050762 Velká_Ledská
18     610086      1031466 Tis
19     615701      1056833 Kostelecká_Lhota
20     660644      1127523 Rytířsko
21     693489      1145417 Býkovec
22     654169      1134485 Kamenička
23     654795      1131408 Řehořov
24     658204      1136186 Vržanov
25     662189      1132137 Otín_nad_Jihlavou
26     662619      1133463 Předboř_nad_Jihlavou
27     662547      1135368 Svatoslav_nad_Jihlavou
28     689139      1152322 Praskolesy_u_Mrákotína
29     675460      1145652 Bezděkov_u_Třešti
30     675068      1144684 Stajiště
31     656458      1119294 Hrbov
32     654096      1122616 Janovice_u_Polné
33     661433      1119351 Nové_Dvory_u_Kamenné
34     827224      1100779 Nedaničky
35     823765      1102625 Petrovice_u_Měčina
36     821938      1099030 Radkovice_u_Měčina
37     828548      1100149 Třebýcina
38     824267      1150283 Filipova_Huť
39     829567      1148543 Javoří_Píla
40     832228      1152889 Roklanský_Les

```

Formát znázorněný na Obrázku 2 však není vhodný k importu dat do databáze, proto bylo nutno vytvořit program, který by tento soubor zpracoval a vytvořil výstup v syntaxi SQL k naplnění DB, respektive sloupců `id`, `nazev_cz`, `x_puv`, `y_puv` v tabulce `SIDLA`. Tyto programy jsou uloženy ve struktuře přiloženého CD znázorněného v Příloze B (`AHajek_BP_2009\Práce_s_tabulkou_sidla\sidla\Sidlo.java` a `Rozdeleni.java`).

Dle tabulky `NALEZITOST` byl naplněn i sloupec `zmeneno` a to na implicitní hodnotu 2, příslušející v této fázi všem sídlům, tzn. všem entitám tabulky `SIDLA`.

Výstupem z programů pozměňujících obsah souborů pro aktualizaci je soubor `sidla-insert.sql`, který je zobrazen na Obrázku 3. Takto zformátovaný vstup byl pomocí rozhraní programu `pgAdmin` nahrán do struktury DB a ukázka tohoto naplnění je na Obrázku 4. (Pozn.: sloupec `nazev_n` zde vyobrazený byl později přesunut do tabulky `SIDLA_MULLER` z důvodu konzistentnosti dat získaných z Müllerových map, kde se právě německý název sídel vyskytuje).

Obrázek 3: Náhled souboru sidla-insert.sql vytvořený pro naplnění databáze prvotními daty

```

INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (1, 'Rájec', 1063020.0, 614503.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (2, 'Čičová', 1060460.0, 619751.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (3, 'velkáČermná_nad_Orlicí', 1059198.0, 620463.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (4, 'Domašín_u_Černíkovíc', 1048871.0, 613703.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (5, 'Skršice', 1036572.0, 627791.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (6, 'Jedlová_v_Orlických_horách', 1037438.0, 603618.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (7, 'Hlinné_u_Dobrého', 1041098.0, 609073.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (8, 'Kamenice_u_Dobrého', 1038393.0, 610321.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (9, 'Rovné_u_Dobrého', 1037832.0, 607984.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (10, 'Běstvíný', 1032747.0, 617351.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (11, 'Domašín_u_Dobrušky', 1036437.0, 613919.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (12, 'Křovice', 1034778.0, 616225.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (13, 'Mělčany_u_Dobrušky', 1037702.0, 616723.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (14, 'Pulice', 1035211.0, 618553.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (15, 'Spáleniště', 1037894.0, 611114.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (16, 'Vyhnánov', 1057912.0, 612012.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (17, 'velká_Ledská', 1050762.0, 617015.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (18, 'Tis', 1031466.0, 610086.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (19, 'Kostelecká_Lhota', 1056833.0, 615701.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (20, 'Rytiřsko', 1127523.0, 660644.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (21, 'Bykovec', 1145417.0, 693489.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (22, 'Kamenička', 1134485.0, 654169.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (23, 'Řehořov', 1131408.0, 654795.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (24, 'Vržanov', 1136186.0, 658204.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (25, 'Otín_nad_Jihlavou', 1132137.0, 662189.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (26, 'Předboř_nad_Jihlavou', 1133463.0, 662619.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (27, 'Svatoslav_nad_Jihlavou', 1135368.0, 662547.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (28, 'Praskolesy_u_Mrákotína', 1152322.0, 689139.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (29, 'Bezděkov_u_Třešti', 1145652.0, 675460.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (30, 'Stajístě', 1144684.0, 675068.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (31, 'Hrbov', 1119294.0, 656458.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (32, 'Janovice_u_Polné', 1122616.0, 654096.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (33, 'Nové_Dvory_u_Kameně', 1119351.0, 661433.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (34, 'Nedamíčky', 1100779.0, 827224.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (35, 'Petrovice_u_Měčina', 1102625.0, 823765.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (36, 'Radkoviце_u_Měčina', 1099030.0, 821938.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (37, 'Třebýcina', 1100149.0, 828548.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (38, 'Filipova_Hut', 1150283.0, 824267.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (39, 'Javorí_Píla', 1148543.0, 829567.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (40, 'Roklanský_Les', 1152889.0, 832228.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (41, 'Vchynice-Tetov_II', 1150296.0, 826070.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (42, 'Bystré_u_Klatov', 1112827.0, 827346.0);
INSERT INTO SIDLA (ID, NAZEV_CZ, X_CPUV, Y_PUV) VALUES (43, 'Hoštice_u_Mochtína', 1110649.0, 829602.0);

```

Obrázek 4: Ukázka DB sídel naplněné prvotními daty

| id | nazev_cz | nazev_n | x_puv | y_puv | x_nova | y_nova | zmeneno |
|-------------|-----------------|---------|---------------|---------------|---------------|---------------|---------|
| [PK] bigint | text | text | double precis | double precis | double precis | double precis | integer |
| 1 | Rájec | | 1063020 | 614503 | | | 2 |
| 2 | Čřčová | | 1060460 | 619751 | | | 2 |
| 3 | Velká_Černná_J | | 1059198 | 620463 | | | 2 |
| 4 | Domašín_u_Čerri | | 1048871 | 613703 | | | 2 |
| 5 | Skršice | | 1036572 | 627791 | | | 2 |
| 6 | Jedlová_v_Orlic | | 1037438 | 603618 | | | 2 |
| 7 | Hlinné_u_Dobré | | 1041098 | 609073 | | | 2 |
| 8 | Kamenice_u_Do | | 1038393 | 610321 | | | 2 |
| 9 | Rovné_u_Dobré | | 1037832 | 607984 | | | 2 |
| 10 | Běštviný | | 1032747 | 617351 | | | 2 |
| 11 | Domašín_u_Dob | | 1036437 | 613919 | | | 2 |
| 12 | Křovice | | 1034778 | 616225 | | | 2 |
| 13 | Měličany_u_Dob | | 1037702 | 616723 | | | 2 |

2.2 Aktualizace databáze zpřesněnými souřadnicemi sídel

Dalším krokem bylo importování zpřesněných souřadnic, respektive naplnění sloupců `x_nova`, `y_nova` v tabulce `SIDLA`. K tomu účelu byly vytvořeny soubory `sidlaXaktual.stx` a `sidlaXsmaz.stx` (kde $X = \{1,2,3,4,5\}$) (viz. Obrázek 5 a 6) pro aktualizaci databáze, první zmíněné jsou pro sídla, která byla na mapových listech Müllerových map nalezena, druhé pro sídla, která na mapových listech nalezena nebyla (více tabulka `NALEZITOST`). Zpřesňování souřadnic probíhalo na podkladě základních map `ZM10`, na kterých byly souřadnice jednotlivých sídel odměřovány (více viz.[2]).

Obrázek 5: Část aktualizací souboru `sidla1aktual.stx`

```
sidla;YX;CM;0;0
10;617618.32;1032724.68;Běstviny
18;610177.80;1031491.33;Tis
126;854530.33;1002401.16;Fojtov
127;861551.55;999781.77;Lesík
129;854875.00;998664.00;oldřichov_u_Nejdku
131;858123.01;1000956.14;suchá_u_Nejdku
132;857987.00;997455.00;Tisová_u_Nejdku
133;855517.00;999722.00;vysoká_Štola
134;855567.96;1007521.14;Jimlíkov
135;852483.75;1006020.42;Mezirolí
138;843213.34;1001929.67;Dolní_Žďár_u_Ostrova
139;841336.24;999096.76;Hanušov
140;845928.80;1001052.63;Hluboký
141;843174.21;1000454.36;Horní_Žďár_u_Ostrova
142;845292.00;1002624.00;Kfely_u_Ostrova
143;840905.33;1001649.76;květnová
144;840418.18;1000310.17;Maroltov
146;842184.27;1000891.91;vykmanov_u_Ostrova
159;846579.91;1007449.41;Lesov
160;848750.00;1006520.00;Podlesí_u_Sadova
161;842929.91;1007697.07;Stráň
162;850419.12;1020727.98;Dražov
163;850511.44;1022593.89;Hlinky
165;833483.42;1000431.12;Korunní
166;833236.44;997342.26;Malý_Hrzín
167;836490.08;999145.43;Osvínov
168;836074.55;999751.63;Peklo
207;630024.91;991720.79;Bobr
208;627192.74;990627.63;Černá_voda_u_Žacléře
209;630191.58;994811.17;Prkenný_Důl
210;631900.82;993141.69;Rýchory
211;631175.78;994938.15;Vernířovice
264;709720.26;1006786.51;Liny
265;703451.65;1025934.54;Struhy
266;699971.89;1016198.67;Bojetice
267;698895.54;1014914.61;Holé_vrchy
268;702500.04;1015340.49;Chloumek_u_Mladé_Boleslavi
269;703787.62;1018517.19;Libichov
270;702591.09;1017027.13;Sýčina
271;698956.63;1016655.18;Týnec_u_Dobrovice
272;698640.32;1017234.48;Úherce
273;690528.54;1010436.65;Bechov
```

Obrázek 6: Část aktualizacího souboru sidla1smaz.stx

```

sidla YX CM 0 0
137 842505.00 999108.00 Arnoldov
332 612909.00 1008059.00 Metujka
758 885198.00 1002788.00 Výspa
915 561740.00 1027482.00 Hundorf
972 739062.00 1001520.00 Štětí_II
1612 821768.00 993692.00 Pavlov_u_Vernéřova
1613 821112.00 993468.00 Potočná_u_Vernéřova
1616 826685.00 1000717.00 Velká_Lesná
1619 828074.00 990204.00 Dolina
1620 828140.00 988352.00 Přisečnice
1624 815688.00 985212.00 Menhartice_u_Křimova
1631 815221.00 995374.00 Kralupy_u_Chomutova
1664 661712.00 991168.00 Přívlika
1736 792727.00 990298.00 Čepirohy
1737 799234.00 985795.00 Dřínov_u_Komořan
1739 796076.00 988686.00 Hořany
1741 788769.00 984204.00 Konobříže
1743 791474.00 986281.00 Most_I
1744 788760.00 984800.00 Pařidla
1748 793188.00 987252.00 Souš
1794 620001.00 1025568.00 Domkov
2226 865754.00 996456.00 Ptačí
2915 860003.00 1005567.00 Stará_Chodovská
2929 876207.00 1007430.00 Leopoldovy_Hamry
3338 825566.00 991573.00 Podmílesy
3361 809429.00 975134.00 Gabrielína_Huť
4365 763249.00 970096.00 Dělouš
4374 763497.00 972738.00 všebořice
4418 668242.00 983351.00 Zlatá_Olešnice_Navarovská
4535 877633.00 1007148.00 Smrčí_u_Nového_Kostela
4536 878289.00 1006369.00 Svažec
4552 888409.00 1010389.00 Velký_Rybník_u_Skalné
4915 864259.00 1016819.00 Vítkov_u_Sokolova
4917 868967.00 1021649.00 Bystřina
5342 800945.00 987197.00 Nové_Sedlo_nad_Bílinou
5345 820476.00 990125.00 Úbočí_u_Výsluní
5348 801729.00 983790.00 Kundratice_u_Chomutova
5349 802443.00 984537.00 Podhůří_u_Vysoké_Peče
5419 804205.00 977260.00 Svahová_I
5422 795605.00 981875.00 Čtrnáct_Dvorců
5435 790207.00 982391.00 Růžodol
5441 797574.00 988345.00 Hořešice

```

Struktura souborů uvedených na Obrázcích 5 a 6 opět nevyhovuje potřebám formátu vstupu do DB a tudíž bylo nutno z aktualizacího souborů vyseparovat data a zapsat je ve vhodné struktuře použitelné pro aktualizaci a uložení do DB.

Programy, které toto umožňují, jsou uloženy v příloženém CD (AHajek_BP_2009\Práce_s_tabulkou_sidla\aktualizace_sidla_aktual a aktualizace_sidla_smaz, ve kterých je v každém z nich samostatná dvojice souborů Sidlo.java a Rozdeleni.java).

Díky těmto programům, které zpracovaly aktualizacího soubory sidlaXaktual.sql a sidlaXsmaz.sql, se vygeneroval výstup znázorněný na Obrázku 7.

Obrázek 7: Část souboru pro aktualizaci DB sídel vygenerovaný programem ze souborů sidlaXaktual.stx

```
UPDATE sidla SET zmeneno = 1, x_nova = 1032724.68, y_nova = 617618.32
WHERE id = 10;
UPDATE sidla SET zmeneno = 1, x_nova = 1031491.33, y_nova = 610177.8
WHERE id = 18;
UPDATE sidla SET zmeneno = 1, x_nova = 1002401.16, y_nova = 854530.33
WHERE id = 126;
UPDATE sidla SET zmeneno = 1, x_nova = 999781.77, y_nova = 861551.55
WHERE id = 127;
UPDATE sidla SET zmeneno = 1, x_nova = 998664.0, y_nova = 854875.0
WHERE id = 129;
UPDATE sidla SET zmeneno = 1, x_nova = 1000956.14, y_nova = 858123.01
WHERE id = 131;
UPDATE sidla SET zmeneno = 1, x_nova = 997455.0, y_nova = 857987.0
WHERE id = 132;
UPDATE sidla SET zmeneno = 1, x_nova = 999722.0, y_nova = 855517.0
WHERE id = 133;
UPDATE sidla SET zmeneno = 1, x_nova = 1007521.14, y_nova = 855567.96
WHERE id = 134;
```

Pozn.: Výstup ze souborů sidlaXsmaz.stx je obdobný situaci na Obrázku 7, jen s tím rozdílem, že hodnota zmeneno se nastavuje na 0.

Celkovým výstupem z této části aktualizace bylo 10 souborů, které se nahrály do databáze a označení ve sloupci zmeneno bylo nahrazeno z implicitní hodnoty 2 na 1 pro sídla v souborech sidlaXaktual.stx a hodnotou 0 pro sídla ze souborů sidlaXsmaz.stx (viz. tabulka NALEZITOST).

V této fázi byla tabulka SIDLA naplněna počtem 22215 záznamů, což neznačí konečný počet sídel. V DB byla v této fázi zanesena pouze sídla poskytnutá ČSÚ a ta byla následně aktualizována. Ukázka takto naplněné tabulky SIDLA je vyobrazena na Obrázku 8.

Pro určení správnosti vložení a aktualizace prvotních dat jsem provedl kontrolní analýzu Příkazem 1 zahrnující vypsání všech sídel majících ve sloupci zmeneno hodnotu 2. Předpoklad byl, že by všechny záznamy v tabulce SIDLA měly mít buď hodnotu 1 nebo 0. Výsledek je zachycen na Obrázku 9.

Obrázek 8: Výpis z tabulky SIDLA po aktualizaci ze souborů sidlaXaktual.stx

| | id | nazev_cz text | nazev_n text | x_puv double precis | y_puv double precis | x_nova double precis | y_nova double precis | zmeneno integer |
|------------|-----------|-------------------------|------------------------|----------------------------------|----------------------------------|-----------------------------------|-----------------------------------|---------------------------|
| 154 | 154 | Kolešov_u_Žlutih | | 1032495 | 829652 | | | 2 |
| 155 | 155 | Močidlec | | 1034118 | 828080 | | | 2 |
| 156 | 156 | Novosedly_u_Ži | | 1035767 | 829853 | | | 2 |
| 157 | 157 | Semtěš_u_Žlutih | | 1031606 | 833432 | | | 2 |
| 158 | 158 | Bor_u_Karlovyčl | | 1007922 | 844856 | 1007954.67 | 844890.85 | 1 |
| 159 | 159 | Lesov | | 1007434 | 846607 | 1007449.41 | 846579.91 | 1 |
| 160 | 160 | Podleší_u_Sado | | 1006520 | 848750 | 1006520 | 848750 | 1 |
| 161 | 161 | Stráň | | 1007666 | 842938 | 1007697.07 | 842929.91 | 1 |
| 162 | 162 | Dražov | | 1020719 | 850354 | 1020727.98 | 850419.12 | 1 |
| 163 | 163 | Hlínky | | 1022692 | 850718 | 1022593.89 | 850511.44 | 1 |
| 164 | 164 | Boč | | 998964 | 832861 | 998828.45 | 832864.81 | 1 |
| 165 | 165 | Korunní | | 1000425 | 833493 | 1000431.12 | 833483.42 | 1 |
| 166 | 166 | Malý_Hrzní | | 997350 | 833244 | 997342.26 | 833236.44 | 1 |

Příkaz 1: Kontrola vložených dat do tabulky SIDLA

```

SELECT
    id,
    nazev_cz
FROM
    sidla
WHERE
    zmeneno = 2
;

```

Obrázek 9: Vyhledání zbylých sídel s atributem zmeneno = 2 v kompletně aktualizované tabulce SIDLA

| | id bigint | nazev_cz text |
|----------|----------------------|--------------------------|
| 1 | 4399 | Vrbka_u_Opavy |
| 2 | 10023 | Košátky |
| 3 | 14740 | Bílá_Bříza |
| 4 | 14741 | Svoboda |
| 5 | 15148 | Ryžovna |
| 6 | 16424 | Služovice |

Jak je z Obrázku 9 patrné, v aktualizacích souborů se nenacházelo 6 sídel. Vzhledem k tomu, že se jedná o velmi malý podíl z celku, nebyly prozatím uskutečněny kroky k nápravě tohoto stavu.

2.3 Uložení dat z Müllerových map

Další krok již vedl k vytvoření nové tabulky ve struktuře DB a to tabulky SIDLA_MULLER. Tato tabulka slouží pro uložení dat získaných pouze ze zpracovaných rukopisných Müllerových map. Jedná se o mapy, na kterých jsou znázorněny kraje Berounský, Prácheňský, Litoměřický, Rakovnický, Chudimský a Hradecký (viz. Příloha A). Předpokládá se využití této DB i v pozdějších projektech, respektive plánuje se začlenění dat z dalších mapových listů rukopisných Müllerových map Čech, ale i z jiných mapových děl, a to nejen pro, v této práci použité, retrospektivní analýzy změn sídel v průběhu dějin.

K této tabulce přibýly ještě dvě referenční a to KRAJ a TYP_SIDELE. Tabulka KRAJ slouží k uchování názvů a označení krajů a TYP_SIDELE pro uchování druhu sídel dle kategorií používaných na počátku 18. století. Důvodem pro vytvoření těchto dvou tabulek bylo normalizování datového modelu a zabránění ukládání chybných či duplicitních hodnot ve sloupcích kraj a typ v tabulce SIDLA_MULLER. Souřadnice, které byly vloženy do této tabulky, se získaly transformací z místní souřadnicové soustavy mapových listů do systému JTSK a tyto ztransformované souřadnice byly použity do sloupců x a y v této tabulce (více v [2], [3]).

Při získávání dat z map byl zjištěn fakt, že na nich byla nalezena i sídla, která se v tabulce SIDLA nevyskytují. Tato sídla byla samozřejmě poznamenána a to s id 10000000 zvyšujícím se po jedničce. Vzhledem k vazbě mezi sloupcem id_sidla tabulky SIDLA_MULLER a id z tabulky SIDLA, bylo nutno sídla nevyskytující se v tabulce SIDLA do ní doplnit.

Poskytnuté údaje pro naplnění tabulky byly ve formě textových souborů, které měly následující strukturu: na každém řádku (kromě hlavičky) bylo uvedeno vygenerované id z tabulky SIDLA nebo nově vytvořené osmimístné, souřadnice Y, souřadnice X (oboje souřadnice v systému Jednotné trigonometrické sítě katastrální), hodnota ve sloupci nalezí, značící příslušnost ke kraji na daném mapovém listě a název sídla v německém jazyce (u kterého byla mezera v názvu reprezentovaná podtržítkem) (více viz. Obrázek 10).

Obrázek 10: Ukázka souboru k naplnění tabulky SIDLA_MULLER

```

Vsechna YX CM 0 0
680 741632.27 1100357.16 0.00 Prtšchitz
2816 780345.81 1082559.80 0.00 Birkenberg
4720 747172.10 1055821.77 0.00 Sbraslaw
6445 760493.55 1094480.07 0.00 Krašna_hora
6454 776475.31 1089027.66 0.00 Milin
6466 756458.00 1074746.00 0.00 Knin
7075 792246.75 1091760.34 1.00 Rosmital
7127 791874.29 1060580.59 0.00 Zbirow
7338 761755.95 1089892.86 0.00 Kameik
7342 745087.38 1091909.45 0.00 Amšchelberg_B.Kozowy_hory
8235 751875.93 1094236.06 0.00 Chlumetz
8839 744801.42 1110107.18 1.00 Gištebnitz
9159 747233.75 1069051.40 0.00 Stiechowitz
9188 765272.36 1074172.13 0.00 Dobříž
9367 778470.85 1063956.32 0.00 Lochowitz
9408 788172.00 1063315.00 0.00 Čzerhowitz
9412 784033.25 1065334.83 0.00 Horžowitz
9413 773551.00 1068151.00 0.00 Hoštowitz
10050 778254.07 1101655.81 1.00 Mirowitz
10380 739185.72 1109261.63 1.00 Borotin
10542 738893.43 1086060.98 0.00 Maršchowitz
10933 742649.00 1100916.00 0.00 Sedletz
16080 796651.74 1068394.99 0.00 Mauth
16089 804355.00 1060120.00 1.00 Radnitz
16145 758235.29 1065634.04 0.00 Mnišchek
16163 747484.84 1065164.62 0.00 Dawle
16519 739450.18 1074652.09 0.00 Nedworžitz
16521 740120.69 1081265.19 0.00 Neweklow
16533 726921.32 1068230.55 1.00 Pišcheli
16853 730608.92 1084101.80 1.00 Bištritz
10000219 762248.68 1057581.52 0.00 wudnian
16787 783523.13 1061001.30 0.00 schebrak
9614 778649.82 1082513.53 0.00 Pržibram
9620 784948.75 1096737.14 1.00 Brzesnitz
16130 740591.00 1064873.00 1.00 Eule
7078 749012.91 1090267.95 0.00 seltšchan
9837 733436.12 1094150.05 0.00 wotitz
16512 733678.58 1101203.53 1.00 Milčzin

```

V těchto souborech pro naplnění tabulky SIDLA_MULLER byl samozřejmě předpokládán výskyt duplicitních hodnot id sídel, díky přesahu použitých mapových listů. Avšak pro vložení osmimístných id do tabulky SIDLA bylo nutné každé toto id vložit jen jednou, kvůli jejich jednoznačnosti, aby mohla sloužit jako primární klíč v tabulce SIDLA a do tabulky SIDLA_MULLER naopak vložit všechna.

Z každého poskytnutého souboru pro naplnění jsem vytvořil vždy dva soubory, soubor XXX_sidla_vsechna, kde XXX značí zkratku kraje, obsahující pouze osmimístné id a soubor XXX_sidla_muller, naplňující tabulku SIDLA_MULLER (Obrázek 11).

Obrázek 11: Ukázka souboru pro vkládání záznamů do tabulky SIDLA_MULLER

```
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),680, 'Prtšchitz', 'BER', 1100357.16, 741632.27, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),2816, 'Birkenberg', 'BER', 1082559.8, 780345.81, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),4720, 'Sbraslaw', 'BER', 1055821.77, 747172.1, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),6445, 'Krašna_hora', 'BER', 1094480.07, 760493.55, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),6454, 'Mílin', 'BER', 1089027.66, 776475.31, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),6466, 'Knín', 'BER', 1074746.0, 756458.0, 0) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),7075, 'Rosmital', 'BER', 1091760.34, 792246.75, 1) ;
INSERT INTO sidla_muller (ID, ID_SIDLA, NAZEV_N, KRAJ, X, Y, NALEZI ) VALUES
(NEXTVAL('seq_sidla_muller'),7127, 'Zbirow', 'BER', 1060580.59, 791874.29, 0) ;
```

Pro generování hodnoty sloupce id do tabulky SIDLA_MULLER jsem vytvořil sekvenci, uvedenou v Příkazu 2, která přiřazuje identifikátor ve sloupci id sídlu dle řádku, ve kterém je vloženo.

Příkaz 2: Sekvence pro generování hodnoty id pro tabulku SIDLA_MULLER

```
CREATE SEQUENCE
  seq_sidla_muller
START WITH 1 INCREMENT BY 1;
```

Z jednotlivých souborů XXX_sidla_vsechna jsem sestavil jeden soubor vsechna_komplet.sql, obsahující všechny záznamy osmimístně označených sídel. Z nich jsem vybral všechna sídla po jednom záznamu, tzn. odstranil jsem duplicity a vložil jsem je do tabulky SIDLA (viz. Obrázek 12). Tato sídla jsou označena podle klíče a to takového, že sídla s osmimístním id ležící v ČR začínají číslicí 1, sídla mimo ČR jiným, konkrétně pro Německo je to číslice 9, pro Polsko 4.

Obrázek 12: Tabulka SIDLA obsahující osmimístná id

| | id | nazev_cz | x_puv | y_puv | x_nova | y_nova | zmeneno | typ |
|-------|-------------|-----------------|---------------|---------------|---------------|---------------|---------|------|
| | [PK] bigint | text | double precis | double precis | double precis | double precis | integer | text |
| 22212 | 22212 | Tyršův_vrch | 1046600 | 741200 | | | 0 | |
| 22213 | 22213 | Pančrác-průmysl | 1047200 | 741700 | | | 0 | |
| 22214 | 22214 | Reitknechtka | 1047000 | 741700 | | | 0 | |
| 22215 | 22215 | Braník-jih_B | 1049400 | 744600 | | | 0 | |
| 22216 | 10000001 | | | | | | 3 | |
| 22217 | 10000002 | | | | | | 3 | |
| 22218 | 10000003 | | | | | | 3 | |
| 22219 | 10000004 | | | | | | 3 | |
| 22220 | 10000005 | | | | | | 3 | |
| 22221 | 10000006 | | | | | | 3 | |
| 22222 | 10000007 | | | | | | 3 | |
| 22223 | 10000008 | | | | | | 3 | |
| 22224 | 10000009 | | | | | | 3 | |

Pozn.: Počet řádků znázorněný na Obrázku 12 pro tabulku SIDLA nekorresponduje s konečným počtem zpracovaných záznamů sídel. Ten je vyšší, než byl v momentě vložení obrázku. Vzhled tabulky SIDLA_MULLER po vložení všech osmimístných id je zobrazen na Obrázku 13.

Obrázek 13: Vyobrazení tabulky SIDLA_MULLER po aktualizaci bez udání typu sídel

| | id [PK] integer | id_sidla integer | nazev_n text | kraj text | x double precis | y double precis | nalezi integer | typ integer |
|----|--------------------|---------------------|-----------------|--------------|--------------------|--------------------|-------------------|----------------|
| 1 | 1 | 680 | Průčhitz | BER | 1100357.16 | 741632.27 | 0 | |
| 2 | 2 | 2816 | Birkenberg | BER | 1082559.8 | 780345.81 | 0 | |
| 3 | 3 | 4720 | Sbraslaw | BER | 1055821.77 | 747172.1 | 0 | |
| 4 | 4 | 6445 | Krařna_hora | BER | 1094480.07 | 760493.55 | 0 | |
| 5 | 5 | 6454 | Millin | BER | 1089027.66 | 776475.31 | 0 | |
| 6 | 6 | 6466 | Knin | BER | 1074746 | 756458 | 0 | |
| 7 | 7 | 7075 | Rosmital | BER | 1091760.34 | 792246.75 | 1 | |
| 8 | 8 | 7127 | Zbirow | BER | 1060580.59 | 791874.29 | 0 | |
| 9 | 9 | 7338 | Kameik | BER | 1089892.86 | 761755.95 | 0 | |
| 10 | 10 | 7342 | Amšchelberg_B | BER | 1091909.45 | 745087.38 | 0 | |
| 11 | 11 | 8235 | Chlumetz | BER | 1094236.06 | 751875.93 | 0 | |
| 12 | 12 | 8839 | Giftebnitz | BER | 1110107.18 | 744801.42 | 1 | |
| 13 | 13 | 9159 | Stiechowitz | BER | 1069051.4 | 747233.75 | 0 | |

Dalším krokem bylo vložení jednotlivých typů sídel do sloupce typ. Proto byly vytvořeny soubory, které obsahují sídla stejného typu v kraji (viz. Obrázek 14).

Obrázek 14: Soubor se sídly rozdělenými dle typu (zde vsi se zámky)

| PRA_vsisezamky | YX | CM | 0 | 0 | |
|----------------|---------|---------|------|---|------------------------|
| 173 | 5719.61 | 3848.87 | 0.00 | | Schikowi |
| 205 | 4489.36 | 4023.84 | 0.00 | | Matßchitz |
| 502 | 5852.57 | 4095.62 | 0.00 | | Köllerdorf_B.Girßitzna |
| 538 | 6258.89 | 3565.34 | 0.00 | | Nemlkau |
| 1407 | 3461.93 | 3527.13 | 0.00 | | Strahl_B.Střelna |
| 2092 | 4163.05 | 4615.18 | 0.00 | | Přetßchín |
| 2160 | 1125.35 | 3065.29 | 0.00 | | Kržeßstiwitz |
| 2175 | 6034.72 | 3727.90 | 0.00 | | Pržeßtanitz |
| 2188 | 5457.65 | 3361.62 | 0.00 | | Czegkow |
| 2191 | 5424.16 | 3568.97 | 0.00 | | Tedraßchitz |
| 2196 | 4394.32 | 2434.70 | 0.00 | | Laßchan |
| 2577 | 6119.24 | 3275.24 | 1.00 | | Gindřichowitz |
| 2580 | 6025.14 | 3136.30 | 0.00 | | Mlaßow |
| 2586 | 5955.94 | 3401.55 | 0.00 | | Autßchín |
| 3451 | 2211.97 | 1205.54 | 0.00 | | Toußchkow |
| 3859 | 3798.12 | 2358.09 | 0.00 | | Polle |
| 4322 | 5964.82 | 4262.32 | 0.00 | | Ober_Teßchau |
| 4327 | 5858.23 | 4453.95 | 0.00 | | Kundratitz |
| 4334 | 5578.49 | 4460.27 | 0.00 | | Wadatitz |
| 5514 | 4075.74 | 3908.23 | 0.00 | | Wohrazenitz |
| 5517 | 3255.64 | 4808.63 | 0.00 | | Cžernetitz |
| 5557 | 4034.61 | 2346.07 | 0.00 | | Bezdiekow |
| 7291 | 2713.02 | 5263.31 | 0.00 | | Budkowa |
| 7310 | 3603.89 | 4958.53 | 0.00 | | Eltßchowitz |
| 7545 | 2306.87 | 1805.12 | 0.00 | | Rakowitz |
| 7554 | 1762.48 | 2197.92 | 0.00 | | Warwaßchau |
| 7625 | 4393.87 | 2137.48 | 0.00 | | woßseletz |
| 7824 | 4493.22 | 5144.95 | 0.00 | | Stíkau |
| 7834 | 3517.90 | 2714.26 | 0.00 | | Bratronitz |
| 7842 | 3840.73 | 4439.40 | 0.00 | | Czeßchtitz |
| 7845 | 2283.32 | 4226.55 | 0.00 | | Drahenitz |
| 8030 | 2282.18 | 2334.25 | 0.00 | | Cžerhonitz |
| 8032 | 2108.03 | 2842.16 | 0.00 | | Cžißchowa |
| 8044 | 2205.61 | 3560.83 | 0.00 | | Keßtržan |
| 8237 | 2148.99 | 795.05 | 0.00 | | Zbenitz |
| 8825 | 4742.32 | 3655.04 | 0.00 | | Schihowitz |
| 9091 | 5418.51 | 4156.34 | 0.00 | | Langendorf |
| 9100 | 6142.78 | 3860.07 | 0.00 | | Hlawniowitz |
| 9104 | 5526.08 | 3603.22 | 0.00 | | Hradek |
| 9106 | 4166.09 | 2430.69 | 0.00 | | Kanowitz |
| 9190 | 2828.91 | 1480.51 | 0.00 | | Drahenitz |
| 9423 | 3219.30 | 4296.81 | 0.00 | | Hoßtitz |
| 9443 | 1659.79 | 4965.73 | 0.00 | | Libietitz |

Formát vstupu na Obrázku 14 byl zpracován do vhodné podoby programy, které jsou uloženy na příloženém CD (\Hajek_BP_2009\Kraj_XXX_YYY, kde v každém ze souborů obsahující aktualizací data v rámci krajů, jsou dvojice programů Sidlo.java a Rozdeleni.java, které s těmito daty pracují).

Výsledkem výstupu programů byl soubor pro vložení typu sídla k příslušným prvkům v tabulce SIDLA_MULLER. Ukázka formátu dat pro vložení typu sídel je vyobrazena na Obrázku 15.

Obrázek 15: Ukázka příkazů, díky kterým se vloží do tabulky SIDLA_MULLER informace o typu sídla

```
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 515 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 3846 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 5825 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 7296 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 8823 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 8824 AND kraj = 'PRA' ;
UPDATE sidla_muller SET typ = 8 WHERE id_sidla = 10326 AND kraj = 'PRA' ;
```

Data ze všech poskytnutých krajů byla zpracována a v této fázi byla tabulka SIDLA_MULLER již plně aktualizovaná. Její částečný přehled naleznete na Obrázku 16.

Obrázek 16: Ukázka plně naplněné tabulky SIDLA_MULLER

| id | id [PK] | id_sidla | nazev_n | kraj | x | y | nalezi | typ |
|-----|---------|----------|------------------------|------|---------------|---------------|---------|---------|
| | integer | integer | text | text | double precis | double precis | integer | integer |
| 930 | 930 | 16134 | Garlik | BER | 1057937.93 | 754126.32 | 0 | 7 |
| 931 | 931 | 16159 | Ober_Cžernulitz | BER | 1056051 | 751952 | 0 | 7 |
| 932 | 932 | 16780 | Fraßich | BER | 1051587 | 765152 | 0 | 7 |
| 933 | 933 | 16786 | Zditz | BER | 1057899 | 777031 | 0 | 7 |
| 934 | 934 | 16817 | Drahno_Aujeßt | BER | 1059627.67 | 795323.57 | 0 | 7 |
| 935 | 935 | 16832 | Rot_Aujeßt | BER | 1103872.72 | 737176.53 | 0 | 7 |
| 936 | 936 | 20399 | Sliwetz | BER | 1087926.6 | 777211.7 | 0 | 7 |
| 937 | 937 | 10064 | Protiwin | PRA | 1139028.09 | 771376.84 | 0 | 1 |
| 938 | 938 | 16879 | Weißich_Birken_B.Wilab | PRA | 1149150 | 791544 | 0 | 1 |
| 939 | 939 | 536 | Unf:_Fr:_Bergftädtl | PRA | 1122987 | 826207 | 0 | 2 |
| 940 | 940 | 7075 | Roßenthal_B.Rozmital | PRA | 1091457.54 | 790080.4 | 0 | 2 |

3. Aplikace metod

3.1 Polohově závislé

Pro polohově závislé analýzy jsem vycházel ze zpřesněných souřadnic v tabulce SIDLA (sloupce *x_nova*, *y_nova*) a z transformovaných souřadnic sídel zobrazených na rukopisných Müllerových mapách, reprezentovanými sloupci *x* a *y* v tabulce SIDLA_MULLER.

3.1.1 Rozdíly souřadnic mezi entitami v tabulce SIDLA a SIDLA_MULLER

První zájmová analýza vychází z relace mezi zmíněnými tabulkami, tzn. že zjišťuji rozdíl mezi souřadnicemi sídla v tabulce SIDLA a souřadnicemi uvedenými v tabulce SIDLA_MULLER. V případě, že se sídlo vyskytuje v tabulce SIDLA_MULLER několikrát, bere se rozdíl pro každý záznam a z nich se následně informativně vypočte jejich průměr.

Již z názvu kapitoly vyplývá, že se jedná o porovnání určení polohy sídel ze dvou různých zdrojů ve dvou odlišných časových obdobích a to samo o sobě již uvozuje domněnku, že k určitým poměrně větším odchylkám dospějí.

Postup byl následující - prvotně jsem sestavil tabulku ROZDIL_SIDLA_MULLER, ze které budou následně odvozené další tabulky, SQL dotazem uvedeným v Příkazu 3.

Příkaz 3: Sql dotaz pro vytvoření tabulky ROZDIL_SIDLA_MULLER

```
CREATE TABLE rozdil_sidla_muller AS (  
  SELECT  
    sidla.id AS SIDLA_ID,  
    sidla.nazev_cz AS NAZEV_CZ,  
    sidla_muller.id AS MULLER_ID,  
    (abs(sidla.x_nova - sidla_muller.x)) AS DELTA_X,  
    (abs(sidla.y_nova - sidla_muller.y)) AS DELTA_Y  
  FROM  
    sidla,  
    sidla_muller  
  WHERE sidla.id = sidla_muller.id_sidla AND  
    sidla_muller.id_sidla < 1000000  
);
```

Její popis naleznete v Tabulce 7.

Tabulka 7: Popis tabulky ROZDIL_SIDLA_MULLER

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| sidla_id | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| nazev_cz | text | český název sídla |
| muller_id | integer | id entity uvedený ve sloupci id tabulky SIDLA_MULLER |
| delta_x | double precision | rozdíl souřadnic X sobě si odpovídajících sídel |
| delta_y | double precision | rozdíl souřadnic Y sobě si odpovídajících sídel |

Ukázka tabulky ROZDIL_SIDLA_MULLER je zobrazena na Obrázku 17.

Obrázek 17: Ukázka výstupu z tabulky ROZDIL_SIDLA_MULLER

| | sidla_id bigint | nazev_cz text | muller_id integer | delta_x double precis | delta_y double precis |
|----|--------------------|------------------|----------------------|--------------------------|--------------------------|
| 1 | 7883 | Ostřetín | 5883 | 494.1499999999 | 671.6999999999 |
| 2 | 7883 | Ostřetín | 4812 | 2086.0700000000 | 1571.0799999999 |
| 3 | 7886 | Poběžovice_u_F | 5885 | 0 | 0 |
| 4 | 16487 | Přelovice | 6250 | 146.8500000000 | 371.7399999999 |
| 5 | 16495 | Řečany_nad_La | 6256 | 0 | 0 |
| 6 | 16503 | Srch | 6262 | 358 | 18.239999999999 |
| 7 | 16506 | Staré_Jesenčan | 6265 | 406.2899999999 | 354.1900000000 |
| 8 | 16507 | Staré_Ždánice | 6266 | 24.360000000001 | 248.8000000000 |
| 9 | 8926 | Veliny | 5966 | 564.8500000000 | 685.4599999999 |
| 10 | 8926 | Veliny | 4951 | 2157.3999999999 | 520.2700000000 |
| 11 | 8936 | Banín | 5975 | 0 | 0 |

5512 rows.

Počet záznamů v této tabulce je 5512, což odpovídá počtu záznamů v tabulce SIDLA_MULLER (6278) po odečtení sídel s 8-místnými souřadnicemi (766) avšak se zanecháním duplicit. U těch záznamů duplicitně uložených sídel, kde je `delta_x`, `delta_y` rovna 0, se jedná o sídla vzata jako identická při transformaci souřadnic z Müllerových map do S-JTSK. Těchto sídel je celkem 1680, pro kraje obsažené v databázi (toto číslo je však ve skutečnosti menší, právě díky obsahu duplicitních hodnot).

Určením polohové odchylky pro každé sídlo nalezené na Müllerových mapách lze udělat jednoduchou statistiku rozložení velikosti a počtu těchto odchylek.

Provedu statistiku rozdělení polohových odchylek v rámci všech, tzn. i duplicitních sídel. Rozdělím tuto úlohu na dvě části, které samostatně vyřeším, pro sídla uložená unikátně a pro sídla uložená duplicitně.

Prvním krokem bude určení odchylek mezi entitami v tabulce SIDLA a entitami v tabulce SIDLA_MULLER s kardinalitou 1 ku 1, tzn. že jednomu záznamu v tabulce SIDLA odpovídá pouze jeden záznam v tabulce SIDLA_MULLER.

Provedu výběr z tabulky ROZDIL_SIDLA_MULLER způsobem uvedeným v Příkazu 4.

Příkaz 4: Výběr unikátních sídel z tabulky ROZDIL_SIDLA_MULLER

```
CREATE TABLE rozdil_sidla_muller_unikatni AS (  
  SELECT  
    sidla_id,  
    delta_x,  
    delta_y,  
    sqrt((delta_x)^2 + (delta_y)^2) AS odchylka  
  FROM  
    rozdil_sidla_muller  
  WHERE  
    sidla_id IN  
    (  
      SELECT  
        id_sidla  
      FROM  
        sidla_muller  
      GROUP BY (id_sidla)  
      HAVING COUNT (id_sidla) = 1  
    )  
);
```

Popis tabulky ROZDIL_SIDLA_MULLER_UNIKATNI je zobrazen v Tabulce 8. Její vzhled je zobrazen na Obrázku 18.

Tabulka 8: Popis tabulky ROZDIL_SIDLA_MULLER_UNIKATNI

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| sidla_id | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| delta_x | double precision | rozdíl souřadnic X sobě si odpovídajících sídel |
| delta_y | double precision | rozdíl souřadnic Y sobě si odpovídajících sídel |
| odchylka | double precision | polohová odchylka mezi záznamy v tabulce SIDLA a jejími odpovídajícími, unikátně uloženými, záznamy v tabulce SIDLA_MULLER |

Obrázek 18: Část tabulky ROZDIL_SIDLA_MULLER_UNIKATNI

| | sidla_id bigint | delta_x double precis | delta_y double precis | odchylka double precis |
|----|--------------------|--------------------------|--------------------------|---------------------------|
| 1 | 7886 | 0 | 0 | 0 |
| 2 | 16487 | 146.850000000 | 371.739999999 | 399.694320825 |
| 3 | 16495 | 0 | 0 | 0 |
| 4 | 16503 | 358 | 18.239999999 | 358.464360292 |
| 5 | 16506 | 406.289999999 | 354.190000000 | 539.001039145 |
| 6 | 16507 | 24.360000000 | 248.800000000 | 249.989698987 |
| 7 | 8936 | 0 | 0 | 0 |
| 8 | 10649 | 560.350000000 | 793.260000000 | 971.212412451 |
| 9 | 10653 | 453.030000000 | 64.380000000 | 457.581648779 |
| 10 | 10662 | 0 | 0 | 0 |
| 11 | 10663 | 34.169999999 | 303.010000000 | 304.930564227 |

4700 rows.

Odchylky jednotlivých záznamů v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI jsem vypočetl podle Vzorce 1.

Ve Vzorcí 2 je potom vypočtena průměrná polohová odchylka záznamů tabulky ROZDIL_SIDLA_MULLER_UNIKATNI.

Vzorec 1: Výpočet polohové odchylky Δ_{xy} pro každé sídlo uložené v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI

$$\Delta_{xy} = \sqrt{(\text{delta } x)^2 + (\text{delta } y)^2}$$

Vzorec 2: Výpočet průměrné polohové odchylky Δ pro záznamy v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI

$$\Delta = \frac{1}{(4700-1)} * \sum (\text{odchylka})$$

Jmenovatel ve Vzorcí 2 je (4700 - 1) proto, že určí průměr na omezené množině sídel, ne pro všechna sídla zanesená v databázi.

Průměrná polohová odchylka tedy dle Vzorce 2 vychází následovně:

$$\Delta = 549,31 \text{ m}$$

Obdobně jako tabulka ROZDIL_SIDLA_MULLER_UNIKATNI, byla vytvořena i tabulka ROZDIL_SIDLA_MULLER_DUPLICITNI. Způsob je uveden v Příkazu 5.

Příkaz 5: Výběr duplicitních sídel z tabulky ROZDIL_SIDLA_MULLER

```
CREATE TABLE rozdil_sidla_muller_duplicitni AS (
    SELECT
        sidla_id,
        delta_x,
        delta_y,
        sqrt((delta_x)^2 + (delta_y)^2) AS odchylka,
        muller_id
    FROM rozdil_sidla_muller
    WHERE sidla_id IN
        (
            SELECT
                id_sidla
            FROM
                sidla_muller
            GROUP BY (id_sidla)
            HAVING COUNT (id_sidla) > 1
        )
);
```

Tabulka ROZDIL_SIDLA_MULLER_DUPLICITNI má parametry, které jsou popsány v Tabulce 10. Ukázka tabulky je na Obrázku 19.

Tabulka 9: Popis tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| sidla_id | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| delta_x | double precision | rozdíel souřadnic X sobě si odpovídajících sídel |
| delta_y | double precision | rozdíel souřadnic Y sobě si odpovídajících sídel |
| odchylka | double precision | polohová odchylka mezi záznamy v tabulce SIDLA a jejími odpovídajícími, duplikátně uloženými, záznamy v tabulce SIDLA_MULLER |
| muller_id | integer | id entity uvedený ve sloupci id tabulky SIDLA_MULLER |

Obrázek 19: Ukázka tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI

| | sidla_id bigint | delta_x double precis | delta_y double precis | odchylka double precis | muller_id integer |
|----|--------------------|--------------------------|--------------------------|---------------------------|----------------------|
| 14 | 8301 | 604.330000000 | 615.809999999 | 862.807455345 | 4831 |
| 15 | 7502 | 598.799999999 | 275.479999999 | 659.128720660 | 5851 |
| 16 | 7502 | 555.339999999 | 53.400000000 | 557.901492738 | 4774 |
| 17 | 7513 | 223.929999999 | 792.060000000 | 823.106122258 | 5859 |
| 18 | 7513 | 101.189999999 | 455.800000000 | 466.897265038 | 4779 |
| 19 | 6965 | 734.929999999 | 1010.279999999 | 1249.31492558 | 5813 |
| 20 | 6965 | 195.910000000 | 79.729999999 | 211.512649740 | 4745 |
| 21 | 6890 | 2348.829999999 | 288.230000000 | 2366.44858422 | 3152 |
| 22 | 6890 | 1734.010000000 | 282.180000000 | 1756.81992033 | 2483 |
| 23 | 8532 | 283.75 | 256.760000000 | 382.674483209 | 4856 |
| 24 | 8532 | 0 | 0 | 0 | 5939 |
| 25 | 8535 | 166.479999999 | 292.939999999 | 336.941291622 | 5942 |

812 rows.

Pro sídlo mající v tabulce SIDLA ve sloupci id hodnotu 8532 platí, že má v jednom záznamu rozdíl souřadnic 0, protože bylo bráno jako identické, a ve druhém již ne, což odpovídá tomu, že bylo nalezeno i na sousedním listě, ale zde už nebylo bráno jako identické pro transformaci souřadnic z místní soustavy mapového listu do S-JTSK.

Provedu výpočet polohové odchylky pro každé sídlo uložené v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI. Způsob je zobrazený ve Vzorcí 3.

Vzorec 3: Výpočet polohové odchylky Δ_{xy} pro každé sídlo uložené v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI

$$\Delta_{xy} = \sqrt{(\text{delta } x)^2 + (\text{delta } y)^2}$$

Vzorec 4: Výpočet průměrné polohové odchylky Δ pro záznamy v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI

$$\Delta = \frac{1}{(812-1)} * \sum (\text{odchylka})$$

Jmenovatel zlomku (812-1) je zvolen ze stejného důvodu jako ve Vzorcí 2. Průměrná polohová odchylka tedy dle Vzorce 4 vychází následovně:

$$\Delta = 502,50 \text{ m}$$

Nyní již přejdu ke statistice v rámci zařazení hodnot odchylek v tabulce ROZDIL_SIDLA_MULLER_UNIKATNI do intervalů pomocí jednoduchého SQL dotazu uvedeným v Příkazu 6 a Příkazu 7.

Příkaz 6: Výběr hodnot odchylka z tabulky ROZDIL_SIDLA_MULLER_UNIKATNI dle intervalů

```
SELECT
*
FROM
rozdil_sidla_muller_unikatni
WHERE
odchylka = 0;
```

Apriori jsem zvolil intervaly odchylek po 100 m, od 0 až do 2000, po 1000 m od 2000 m do 4000 m a hodnoty větší než 4000 m (viz. Tabulky 10 a 11).

Tabulka 10: Přehled odchylek tabulky ROZDIL_SIDLA_MULLER_UNIKATNI

| Velikost odchylky | Počet záznamů |
|-------------------|---------------|
| 0 | 1355 |
| 0 – 100 | 64 |
| 100 – 200 | 141 |
| 200 – 300 | 268 |
| 300 – 400 | 350 |
| 400 – 500 | 400 |
| 500 – 600 | 354 |
| 600 – 700 | 310 |
| 700 – 800 | 267 |
| 800 – 900 | 245 |
| 900 – 1000 | 168 |
| 1000 – 1100 | 147 |
| 1100 – 1200 | 133 |
| 1200 – 1300 | 85 |
| 1300 – 1400 | 80 |
| 1400 – 1500 | 56 |
| 1500 – 1600 | 39 |
| 1600 – 1700 | 40 |
| 1700 – 1800 | 35 |
| 1800 – 1900 | 26 |
| 1900 – 2000 | 29 |
| 2000 – 3000 | 71 |
| 3000 – 4000 | 26 |
| > 4000 | 11 |
| maximum | 13075,19 m |

Následně jsem vytvořil statistiku rozdělení hodnot odchylek v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI do intervalů dle SQL dotazu uvedeným v Příkazu 7.

Příkaz 7: Výběr hodnot odchylka z tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI dle intervalů

```
SELECT
    *
FROM
    rozdil_sidla_muller_duplicitni
WHERE
    odchylka = 0;
```

Zvolil jsem stejné rozdělení intervalů jako pro Tabulku 10.

Tabulka 11: Přehled odchylek tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI

| Velikost odchylky | Počet záznamů |
|-------------------|---------------|
| 0 | 325 |
| 0 – 100 | 7 |
| 100 – 200 | 29 |
| 200 – 300 | 47 |
| 300 – 400 | 48 |
| 400 – 500 | 41 |
| 500 – 600 | 52 |
| 600 – 700 | 34 |
| 700 – 800 | 38 |
| 800 – 900 | 30 |
| 900 – 1000 | 28 |
| 1000 – 1100 | 18 |
| 1100 – 1200 | 14 |
| 1200 – 1300 | 16 |
| 1300 – 1400 | 15 |
| 1400 – 1500 | 8 |
| 1500 – 1600 | 10 |
| 1600 – 1700 | 9 |
| 1700 – 1800 | 5 |
| 1800 – 1900 | 8 |
| 1900 – 2000 | 4 |
| 2000 – 3000 | 21 |
| 3000 – 4000 | 4 |
| > 4000 | 1 |
| maximum | 13307,69 m |

Sestavením tabulky ODCHYLKY_SIDLA_MULLER získám tabulku ve které budou zaneseny odchylky z tabulky ROZDIL_SIDLA_MULLER_UNIKATNI a průměrné odchylky z tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI. To se provedlo Příkazem 8.

Příkaz 8: Tvorba tabulky ODCHYLKY_SIDLA_MULLER

```
CREATE TABLE odchylky_sidla_muller AS (
  SELECT DISTINCT ON (a.sidla_id)
    a.sidla_id,
    (a.odchylka + b.odchylka)/2 AS odchylka
  FROM
    rozdil_sidla_muller_duplicitni a,
    rozdil_sidla_muller_duplicitni b,
    sidla s
  WHERE
    a.sidla_id = b.sidla_id AND
    a.muller_id <> b.muller_id AND
    s.id = a.sidla_id
  UNION
  SELECT
    sidla_id,
    odchylka
  FROM
    rozdil_sidla_muller_unikatni
);
```

Vlastní podoba tabulky vychází ze stejné struktury jak tabulky ROZDIL_SIDLA_MULLER_UNIKATNI a ROZDIL_SIDLA_MULLER_DUPLICITNI, z nich však přejímá pouze sloupce sidla_id, odchylka. Ukázka je na Obrázku 20.

Obrázek 20: Částečná ukázka tabulky ODCHYLKY_SIDLA_MULLER

| | sidla_id bigint | odchylna double precis |
|------------|--------------------|---------------------------|
| 1 | 2 | 456.641190599 |
| 2 | 4 | 501.266588653 |
| 3 | 5 | 41.9280156935 |
| 4 | 7 | 576.717081592 |
| 5 | 8 | 0 |
| 6 | 9 | 1321.04302556 |
| 7 | 10 | 1742.57580497 |
| 8 | 11 | 0 |
| 9 | 12 | 389.503784962 |
| 10 | 13 | 678.562900547 |
| 11 | 14 | 608.421468556 |
| 5106 rows. | | |

Počet záznamů 5106 v tabulce ODCHYLKY_SIDLA_MULLER odpovídá počtu existujících sídel na Müllerových mapách s unikátně uloženými duplicitními záznamy. V číslech tomu odpovídá: počet všech záznamů v tabulce SIDLA_MULLER (6278) bez počtu neexistujících sídel (766) a bez polovičního počtu dvojně uložených sídel (406) a tím dojde k číslu 5106.

Samotná statistika byla provedena nad dílčími úlohami pro unikátně a duplicitně uložená sídla, nikoli však pro data uložená v tabulce ODCHYLKY_SIDLA_MULLER proto, aby se z průměrovaných hodnot odchylek duplicitních sídel nezanesla do výpočtu nepřesnost.

Z tabulek 10 a 11 jsem zhotovil grafy znázorňující závislost nenulových velikostí odchylek na jejich četnosti. Grafy jsou v Příloze 1 pro odchylky tabulky ROZDIL_SIDLA_MULLER_UNIKATNI a druhý v Příloze 2 pro odchylky v tabulce ROZDIL_SIDLA_MULLER_DUPLICITNI. Do nich nejsou zaneseny nulové odchylky, pro lepší přehlednost grafu.

Velikost odchylek by se mohla zdát, pro dnešní požadavky kladené na přesnost, neúměrná, avšak vzhledem k metodám měření v době vzniku podkladů se nejedná o zvláště velká čísla. Musíme také počítat s tím, že v průběhu let se mohla měnit i klasifikace sídel, respektive umístění definičních bodů, v rámci přirozeného vývoje sídel.

Průměrné polohové odchylky uvedené ve Vzorcích 2 a 4 reprezentují klasický průměr. O váženém průměru nejde v tomto případě uvažovat, nejen vzhledem k tomu, že při výběru jiných identických bodů při transformaci by se statistický soubor změnil, ale také například neznám průměrné odchylky měření a nejen proto jsem apriori zvolil váhu všech odchylek 1, i když je z grafů jasně patrné, že se větší odchylky vyskytují mnohem méně často než menší. Což by mohlo, díky rozložení velikostí odchylek v grafu, naznačovat možnost použití různých vah.

Co se týká přesných směrů posunutí poloh sídel, jednalo by se o určení vektorů, ale tímto se však dále zabývat nebudu.

Dalším problémem k zamyšlení by byl i fakt, že sídla, která jsou mapována uvnitř mapového listu, to znamená, že náleží kraji, pro který byl mapový list určen (mají ve sloupci nalezi tabulky SIDLA_MULLER hodnotu 0), byla pravděpodobně zaměřena přesněji, než ta, která byla zmapována mimo hranice kraje, ale spadala ještě do oblasti mapového listu. Toto je však domněnka, která by mohla být podrobena dalšímu zkoumání.

3.1.2 Rozdíly souřadnic mezi sídly, která se v tabulce SIDLA_MULLER vyskytují více než jednou

Vzhledem k překryvu mapových listů Müllerových rukopisných map je pravděpodobná možnost výskytu jednoho sídla na více listech. Obecně při výskytu této možnosti je předpokládáno, že záznamy tohoto sídla budou mít různě určené souřadnice, a proto v této kapitole provedu také zprůměrování souřadnic těchto sídel pro možnost propojení tabulek SIDLA a SIDLA_MULLER relací s kardinalitou 1:1.

Tato kapitola si klade za cíl nastínit kvalitu přesahu rukopisných Müllerových map krajů, které se mapovaly jednotlivě, jedná se o mapy ostrovního typu, proto je vůbec tato statistika proveditelná.

Prvním krokem bylo především zjistit počet duplicitních hodnot v tabulce SIDLA_MULLER. Jedná se o sídla znázorněná na území přesahů, a to v databázovém smyslu znamená vyhledání takových různých entit v tabulce SIDLA_MULLER, které mají stejnou hodnotu ve sloupci `id_sidla` tabulky SIDLA_MULLER.

Pro tento účel jsem vytvořil Příkazem 9 tabulku POCET_SHODNYCH_SIDEI.

Příkaz 9: Vytvoření tabulky POCET_SHODNYCH_SIDEI

```
CREATE TABLE pocet_shodnych_sideli AS(  
  SELECT  
    id_sidla,  
    COUNT (id_sidla) AS pocet  
  FROM sidla_muller  
  GROUP BY (id_sidla)  
  HAVING COUNT (id_sidla) > 1  
);
```

Výsledkem byla tabulka, jejíž popis je uveden v Tabulce 12

Tabulka 12: Popis tabulky POCET_SHODNYCH_SIDEI

| Název atributu | Datový typ PostgreSQL | Popis |
|-----------------------|-----------------------|---|
| <code>id_sidla</code> | integer | id entity uvedený ve sloupci <code>id</code> tabulky SIDLA |
| <code>pocet</code> | integer | počet záznamů v tabulce SIDLA_MULLER, které se vztahují k danému záznamu ve sloupci <code>id_sidla</code> |

Část této tabulky je zobrazena na Obrázku 21.

Obrázek 21: Ukázka tabulky POCET_SHODNYCH_SIDEL

| | id_sidla integer | pocet bigint |
|----|---------------------|-----------------|
| 1 | 6890 | 2 |
| 2 | 8929 | 2 |
| 3 | 6424 | 2 |
| 4 | 8305 | 2 |
| 5 | 6441 | 2 |
| 6 | 1500 | 2 |
| 7 | 1825 | 2 |
| 8 | 4183 | 2 |
| 9 | 10910 | 2 |
| 10 | 850 | 2 |
| 11 | 20238 | 2 |
| 12 | 973 | 2 |
| 13 | 3223 | 2 |
| 14 | 15925 | 2 |

413 rows.

Z této statistiky plyne, že počet sídel v tabulce SIDLA, která mají více odpovídajících záznamů v tabulce SIDLA_MULLER je 413 a všechna jsou v ní uvedena dvakrát. Vzhledem ke konfiguraci dat, tzn. dat získaných z tehdejšího Berounského, Prácheňského, Litoměřického, Rakovnického, Chrudimského a Hradeckého kraje, viz. Příloha A a [9] je zřejmé, že žádné sídlo se zde nemůže vyskytovat třikrát

Druhým krokem bylo vytvoření tabulky ROZDIL_V_MULLEROVI vhodným SQL Příkazem 10.

Příkaz 10: Tvorba tabulky ROZDIL_V_MULLEROVI

```
CREATE TABLE rozdil_v_mullerovi AS(
  SELECT
    (abs(a.x - b.x)) AS delta_x,
    (abs(a.y - b.y)) AS delta_y,
    a.nazev_n AS nazev_muller1,
    b.nazev_n AS nazev_muller2,
    s.id AS id_sidla,
    a.id AS id_muller
  FROM sidla_muller a,
       sidla_muller b,
       sidla s
  WHERE a.id_sidla = b.id_sidla AND
        a.id <> b.id AND
        s.id = a.id_sidla
  ORDER BY (a.id_sidla)
);
```

Popis vzniklé tabulky ROZDIL_V_MULLEROVI je uveden v Tabulce 13.

Tabulka 13: Popis tabulky ROZDIL_V_MULLEROVI

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| delta_x | double precision | rozdíl souřadnic X sobě si odpovídajících sídel |
| delta_y | double precision | rozdíl souřadnic Y sobě si odpovídajících sídel |
| nazev_muller1 | text | první název zájmového sídla |
| nazev_muller2 | text | druhý název zájmového sídla |
| id_sidla | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| id_muller | integer | id entity uvedený ve sloupci id tabulky SIDLA_MULLER |

Jak tato tabulka vypadá, je znázorněno na Obrázku 22.

Obrázek 22: Ukázka tabulky ROZDIL_V_MULLEROVI

| | delta_x double precis | delta_y double precis | nazev_muller1 text | nazev_muller2 text | id_sidla bigint | id_muller integer |
|----|--------------------------|--------------------------|-----------------------|-----------------------|--------------------|----------------------|
| 1 | 103.2700000000 | 256.2000000000 | Cziczowa | Cziczowa | 2 | 5540 |
| 2 | 103.2700000000 | 256.2000000000 | Cziczowa | Cziczowa | 2 | 4267 |
| 3 | 636.1999999999 | 373.2399999999 | Schachow | Schachow | 340 | 4313 |
| 4 | 636.1999999999 | 373.2399999999 | Schachow | Schachow | 340 | 5559 |
| 5 | 64.0800000000 | 48.6900000000 | Noinitz | Noinitz | 371 | 2082 |
| 6 | 64.0800000000 | 48.6900000000 | Noinitz | Noinitz | 371 | 3209 |
| 7 | 0 | 0 | Karwatetz | Krawatetz | 373 | 2084 |
| 8 | 0 | 0 | Krawatetz | Karwatetz | 373 | 3210 |
| 9 | 151.4599999999 | 61.75 | Lachowitz | Lachowitz | 375 | 2086 |
| 10 | 151.4599999999 | 61.75 | Lachowitz | Lachowitz | 375 | 3211 |
| 11 | 0 | 0 | Sinutz | Sinutz | 378 | 3212 |

826 rows.

V tabulce ROZDIL_V_MULLEROVI je uvedeno 826 sídel, což odpovídá počtu záznamů v tabulce POCET_SHODNYCH_SIDEL, kde je uvedeno, že každý záznam odpovídající hodnoty ve sloupci id_sidla je dvojí. Z této tabulky následně vychází i analýza zmíněná v kapitole 3.2.2.

Nyní provedu samotnou analýzu statistického rozdělení polohových odchylek duplicitně uložených sídel. V tuto chvíli mi však postačí mít tabulku obsahující duplicitní sídla pouze v jednom záznamu. Budu se zabývat sloupci delta_x, delta_y, id_sidla z tabulky ROZDIL_V_MULLEROVI. Tím pádem vyseparuju z tabulky ROZDIL_V_MULLEROVI vždy jeden ze dvou shodných záznamů a vypočítám jejich odchylku Příkazem 11.

Příkaz 11: Vytvoření tabulky ROZDIL_V_MULLEROVI_UNIKATNE

```
CREATE TABLE rozdil_v_mullerovi_unikatne AS (
  SELECT DISTINCT ON (id_sidla)
    id_sidla,
    delta_x,
    delta_y,
    sqrt((delta_x)^2 + (delta_y)^2) AS odchylka
  FROM
    rozdil_v_mullerovi
);
```

Tabulka ROZDIL_V_MULLEROVI_UNIKATNE je zobrazena na Obrázku 23 a popsána v Tabulce 14.

Tabulka 14: Popis tabulky ROZDIL_V_MULLEROVI_UNIKATNE

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| id_sidla | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| delta_x | double precision | absolutní rozdíl souřadnic X odpovídajících si sídel |
| delta_y | double precision | absolutní rozdíl souřadnic Y odpovídajících si sídel |
| odchylka | double precision | polohová odchylka mezi duplicitně uloženými záznamy v tabulce SIDLA_MULLER |

Obrázek 23: Vyobrazení části tabulky ROZDIL_V_MULLEROVI_UNIKATNE

| | id_sidla bigint | delta_x double precis | delta_y double precis | odchylka double precis |
|----|--------------------|--------------------------|--------------------------|---------------------------|
| 1 | 2 | 103.2700000000 | 256.2000000000 | 276.2302172100 |
| 2 | 340 | 636.1999999999 | 373.2399999999 | 737.6032386040 |
| 3 | 371 | 64.0800000000 | 48.6900000000 | 80.4795781550 |
| 4 | 373 | 0 | 0 | 0 |
| 5 | 375 | 151.4599999999 | 61.75 | 163.5640366950 |
| 6 | 378 | 0 | 0 | 0 |
| 7 | 380 | 21.2000000000 | 107.6899999999 | 109.7568954550 |
| 8 | 474 | 314.9000000000 | 2420.3300000000 | 2440.7292596400 |
| 9 | 477 | 898.3499999999 | 1026.9899999999 | 1364.4563688800 |
| 10 | 620 | 78.4200000000 | 344.75 | 353.5565851450 |
| 11 | 625 | 0 | 0 | 0 |

413 rows.

Data získaná z tabulky ROZDIL_V_MULLEROVI_UNIKATNE jsem rozdělil do intervalů stejným způsobem jako v analýze provedené v Kapitole 3.1.1 a jednoduchým SQL dotazem (obdobně jako Příkazy 6 a 7) jsem vybral ty odchylky, které do jednotlivých intervalů náležely. Výsledky jsou zobrazeny v Tabulce 15.

Tabulka 15: Přehled odchylek tabulky ROZDIL_V_MULLEROVI_UNIKATNE

| Velikost odchylky | Počet záznamů |
|-------------------|---------------|
| 0 | 86 |
| 0 – 100 | 7 |
| 100 – 200 | 21 |
| 200 – 300 | 38 |
| 300 – 400 | 39 |
| 400 – 500 | 31 |
| 500 – 600 | 28 |
| 600 – 700 | 17 |
| 700 – 800 | 27 |
| 800 – 900 | 18 |
| 900 – 1000 | 19 |
| 1000 – 1100 | 13 |
| 1100 – 1200 | 9 |
| 1200 – 1300 | 11 |
| 1300 – 1400 | 11 |
| 1400 – 1500 | 6 |
| 1500 – 1600 | 6 |
| 1600 – 1700 | 5 |
| 1700 – 1800 | 2 |
| 1800 – 1900 | 4 |
| 1900 – 2000 | 1 |
| 2000 – 3000 | 12 |
| 3000 – 4000 | 1 |
| > 4000 | 1 |
| maximum | 13307,69 m |

Grafický přehled údajů z Tabulky 15 je znázorněn v Příloze 3 jako graf přehledu nenulových odchylek na jejich četnosti. Z nich je patrná závislost mezi velikostí odchylky a četností výskytu této odchylky, což odpovídá předpokladu, že velkých odchylek bude znatelně méně než relativně malých. Musíme brát zřetel samozřejmě i na kvalitu mapových podkladů, způsob jejich vyhotovení i na lidský faktor.

Odchylka 0 značí ta sídla, která mají opravdu shodné souřadnice a vzhledem k tomu, že určují odchylky v rámci jednoho druhu mapového podkladu, nejedná se tedy o identická sídla pro transformaci.

Dalším krokem byl výpočet střední hodnoty polohových odchylek způsobem uvedeným ve Vzorcí 5.

Vzorec 5: Výpočet střední polohové odchylky pro data uvedená v tabulce ROZDIL_V_MULLEROVI_UNIKATNE

$$A = \frac{1}{(413-1)} * \sum (\text{odchylka})$$

Číslo 413 značí celkový počet sídel, kterých se statistika týká, odečtení čísla 1 je provedeno ze stejného důvodu jako u Vzorce 2 a 4. Vycházím z předpokladu, že každá odchylka je zatížená stejnou chybou, neznám ani chybu měření, ani přesnosti měřících přístrojů či použitých metod mapování. Avšak vzhledem k době a metodice, kterou byly Müllerovy mapy zhotovovány, nemůžeme vysokou přesnost uvažovat. Tudíž jsem zvolil váhu všech odchylek 1, jako pro případ v kapitole 3.1.1.

Střední polohová odchylka dle Vzorce 5 vyšla:

$$\Delta = 564,08 \text{ m}$$

Analýza odchylek duplicitně uložených sídel v tabulce SIDLA_MULLER byla provedena s tím předpokladem, že veškerá sídla, která byla zmapována na mapových listech mimo příslušný kraj, než pro který byl mapový list určen, byla zaměřena se stejnou přesností v rámci překryvu těchto mapových listů. To znamená předpoklad, že každý přesah mapových listů má stejnou přesnost.

3.2 Polohově nezávislé, atributové

Následující analýzy nebudou operovat s určením sídla na zemském povrchu, to znamená s jejich souřadnicemi, ale soustředím se na jiné atributy, respektive sloupce tabulek, které mají výchozí tabulky SIDLA a SIDLA_MULLER. V tomto případě se jedná o dvě úrovně analýz a to vzhledem k typu sídel (druhé dvě podkapitoly) a k jejich názvům (první dvě podkapitoly).

3.2.1 Úplná shoda názvů sídel

V rámci atributových analýz bylo jednou z možností využití databáze jako podkladu pro porovnání názvu sídla, který je uveden na rukopisných Müllerových mapách (německý) a názvu poskytnutého ČSÚ (český).

Vzhledem k možnostem SQL funkcí a operátorů [7] pro práci s řetězcí byla nakonec zvolena analýza úplné shody názvů sídel, kterou se zde budu nyní zabývat. Co se týká přepisů, přeložení, odvození a dalších způsobů změn názvů sídla, pro analýzu v rámci podkapitoly, ty nebyly brány v potaz. Jednou z možností by bylo porovnávat názvy sídel po znacích, či skupinách znaků, čímž by se vyhledaly například nejčastější překlady hlásek z německého do českého jazyka. Tento a další způsoby porovnávání názvů sídel jsou možné v rámci zahrnutí patřičných metod či zkonstruování vhodného překladače.

Jednoduchým SQL dotazem jsem tedy vybral, pro odpovídající si záznamy v tabulkách SIDLA a SIDLA_MULLER ty záznamy, které mají naprosto shodný název. To je uvedeno v Příkazu 12.

Příkaz 12: Vytvoření tabulky SHODNE_NAZVY

```
CREATE TABLE shodne_nazvy AS (  
  SELECT  
    sidla.id AS SIDLA_ID,  
    sidla.nazev_cz,  
    sidla_muller.id AS MULLER_ID,  
    sidla_muller.nazev_n  
  FROM  
    sidla,  
    sidla_muller  
  WHERE  
    sidla.nazev_cz = sidla_muller.nazev_n AND  
    sidla.id = sidla_muller.id_sidla  
);
```

Tato sídla jsem uložil do nově vzniklé tabulky SHODNE_NAZVY s atributy popsanými v Tabulce 16. Její ukázka je na Obrázku 24.

Tabulka 16: Popis tabulky SHODNE_NAZVY

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|--|
| sidla_id | bigint | id entity uvedený ve sloupci id tabulky SIDLA |
| nazev_cz | text | český název sídla |
| muller_id | integer | id entity uvedený ve sloupci id tabulky SIDLA_MULLER |
| nazev_n | text | německý název sídla |

Obrázek 24: Část tabulky SHODNE_NAZVY

| | sidla_id bigint | nazev_cz text | nazev_n text | muller_id integer |
|-----------|----------------------------|--------------------------|-------------------------|------------------------------|
| 1 | 16789 | Kladno | Kladno | 3137 |
| 2 | 12348 | Chlum | Chlum | 922 |
| 3 | 3910 | Bukol | Bukol | 3375 |
| 4 | 6768 | Slatina | Slatina | 3507 |
| 5 | 7172 | Ruda | Ruda | 3537 |
| 6 | 9176 | Slatina | Slatina | 3589 |
| 7 | 4647 | Krchleby | Krchleby | 835 |
| 8 | 16825 | Chlum | Chlum | 3803 |
| 9 | 10175 | Dlouhopolsko | Dlouhopolsko | 5151 |
| 10 | 5353 | Hrdly | Hrdly | 3933 |
| 11 | 12661 | Chlum | Chlum | 6131 |

101 rows.

Jednoduchým spočtením počtu řádků v této tabulce zjistíme, že u 101 sídel, které jsou obsaženy v tabulce SIDLA_MULLER jsou jejich německé názvy naprosto shodné s jejich odpovídajícími českými v tabulce SIDLA. Jedná se z drtivé většiny o názvy, které neobsahují znaky používané v českém jazyce a to háčky a čárky, které německý jazyk neužívá.

3.2.2 Rozdíl německých názvů sídel v rámci překrývajících se mapových listů

Ze zhotovené tabulky ROZDIL_V_MULLEROVI (více Kapitola 3.1.2) znázorněné na Obrázku 22 je vidět, že se liší názvy sídel nejen mezi českým a německým jazykem, ale i v rámci německého jazyka a to na sousedních mapových listech, respektive na jejich překryvech. Z tohoto faktu tato statistika právě vychází. Přehledně zobrazuje obě kombinace takovýchto sídel. Výběr z tabulky ROZDIL_V_MULLEROVI byl proveden Příkazem 13.

Příkaz 13: Tvorba tabulky VYBER_JINYCH_NEMECKYCH

```
CREATE TABLE vyber_jinych_nemeckych AS(  
  SELECT  
    nazev_muller1,  
    nazev_muller2,  
    id_sidla  
  FROM  
    rozdil_v_mullerovi  
  WHERE  
    nazev_muller1 <> nazev_muller2  
);
```

Struktura tabulky VYBER_JINYCH_NEMECKYCH je tedy odpovídající výběru z tabulky ROZDIL_V_MULLEROVI a je uvedena v Tabulce 17. Její zobrazení je na Obrázku 25.

Tabulka 17: Popis tabulky VYBER_JINYCH_NEMECKYCH

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| nazev_muller1 | text | první název zájmového sídla |
| nazev_muller2 | text | druhý název zájmového sídla |
| id_sidla | bigint | id entity uvedený ve sloupci id tabulky SIDLA |

Obrázek 25: Část přehledu tabulky VYBER_JINYCH_NEMECKYCH

| | nazev_muller1 text | nazev_muller2 text | id_sidla bigint |
|----|-----------------------|-----------------------|--------------------|
| 1 | Sbenitz | Zbenitz | 8237 |
| 2 | Ninßpürg | Ninßpurg | 9064 |
| 3 | Gaßtorf | Gaßtorf_B_Hoska | 6442 |
| 4 | Rzewnitz | Dřewnitz | 9153 |
| 5 | Krawatetz | Karwatetz | 373 |
| 6 | Wehlowitz | Welonitz | 620 |
| 7 | Předwořitz | Prředbořitz | 3425 |
| 8 | Schloukowitz | Sloukowitz | 664 |
| 9 | Klein_Przilep | Klein_Přilep | 826 |
| 10 | Wißcheraud | Wißfehrad | 4681 |
| 11 | Meißlowitz | Meißchlowitz | 477 |

288 rows.

Údaj 288 značí dvojnásobek duplicitních hodnot názvu sídla, protože je v tabulce uveden řádek obsahující název v pořadí nazev_muller1-nazev_muller2 a také i nazev_muller2-nazev_muller1. Takovýchto dvojic je tedy 144, což ze 413 podvojných záznamů stejného sídla činí celých 35%. Toto poměrně vysoké procento je vcelku překvapivé, domnívám se, že by mohlo být zapříčiněno převážně fonetickým přepisem, nelze vyloučit i lidský faktor. Období, v němž se podkladová data získávala, nebylo dlouhé a nepředpokládám, že by v jeho průběhu došlo k přejmenování sídel.

3.2.3 Poměr typů sídel na Müllerových mapách

Další statistikou bylo určení podílu zastoupení sídel v tabulce SIDLA_MULLER v rámci jejich typu. Tento úkol byl rozdělen do čtyř částí, zaprvé určení podílu zastoupení typů sídel pro jednoznačně uložená data v tabulce SIDLA_MULLER, zadruhé pro duplicitně uložená sídla, zatřetí pro všechna sídla v souboru, avšak s jedním záznamem duplicitních hodnot, za čtvrté pro ta sídla, která již neexistují.

Měl jsem v plánu vytvořit tabulku obsahující všechna, i duplicitně uložená data, ale zjistil jsem, že nejednou se v tabulce SIDLA_MULLER vyskytují sídla se dvěma záznamy, ve kterých je uveden jiný typ sídla (příklad uveden na Obrázku 26). Výběr těchto sídel je v Příkazu 14.

Příkaz 14: Výběr rozdílných typů sídel v rámci duplicitních sídel v tabulce SIDLA_MULLER

```
SELECT
  a.id,
  b.id,
  a.typ,
  b.typ,
  a.nazev_n,
  b.nazev_n
FROM
  sidla_muller a,
  sidla_muller b
WHERE
  a.id <> b.id AND
  a.id_sidla = b.id_sidla AND
  a.typ <> b.typ
;
```

Obrázek 26: Ukázka výběru zobrazující rozdílnost typu sídel u duplicit

| | id integer | id integer | typ integer | typ integer | nazev_n text | nazev_n text |
|---|---------------|---------------|----------------|----------------|-----------------|-----------------|
| 1 | 851 | 3951 | 6 | 9 | Ninßpürg | Ninßpurg |
| 2 | 3110 | 2637 | 3 | 2 | Loboßitz | Loboßitz |
| 3 | 3127 | 361 | 2 | 5 | Rzewnitz | Dřewnitz |
| 4 | 857 | 3953 | 6 | 9 | Tachlowitz | Tachlowitz |
| 5 | 858 | 3955 | 6 | 9 | Chrußtenitz | Chrußtenitz |
| 6 | 861 | 3956 | 6 | 9 | Königshof | Königshof |
| 7 | 878 | 3971 | 6 | 9 | Chotetsch | Chotetßch |
| 8 | 924 | 1664 | 7 | 5 | Lažowitz | Laßchowitz |
| - | --- | ---- | - | - | ... | ... |

OK, Unix Ln 14 Col 9 Ch 163 158 rows, 94 ms

Pokud bych však použil tato data k následné statistice, vnesla by se mi do ní chyba právě díky duplicitním sídlům, která mají jiné označení ve sloupci `typ` a to by nebylo přípustné. Neexistuje klíč k tomu, kterou z odlišných hodnot bych měl do statistiky zahrnout a kterou ne. Proto část analýzy (konkrétně Případ 2) s těmito daty nepracuje.

Případ 1: Poměr existujících neredundantních sídel v tabulce SIDLA_MULLER

Nejprve provedu statistiku typu sídel pro taková, která mají v tabulce SIDLA_MULLER pouze jeden záznam a v dnešní době existují, tzn. že mají svoje označení ve sloupci id v tabulce SIDLA menší jak 10000000. K tomu sloužil Příkaz 15.

Příkaz 15: Tvorba tabulky MULLER_UNIQUE_ZAZNAMY

```
CREATE TABLE muller_unique_zaznamy AS(  
  SELECT  
    m.id_sidla,  
    m.typ,  
    t.vyznam  
  FROM  
    sidla_muller m,  
    typ_sidel t  
  WHERE  
    m.id_sidla IN  
    (  
      SELECT  
        id_sidla  
      FROM  
        sidla_muller  
      GROUP BY (id_sidla)  
      HAVING COUNT (id_sidla) = 1  
    )  
    AND m.id_sidla < 10000000  
    AND m.typ = t.typ  
);
```

Tím se vytvořila tabulka MULLER_UNIQUE_ZAZNAMY, která obsahovala existující sídla, unikátně se vyskytující v tabulce SIDLA_MULLER, spolu s určením jejich typu. To je zobrazeno v Tabulce 18 a její ukázka je na Obrázku 27.

Tabulka 18: Popis tabulky MULLER_UNIQUE_ZAZNAMY

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| id_sidla | integer | id sídel uvedených v tabulce SIDLA |
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEL |
| vyznam | text | význam uvedeného typu sídla |

Obrázek 27: Část tabulky MULLER_UNIQUE_ZAZNAMY

| | id_sidla integer | typ integer | vyznam text |
|----|---------------------|----------------|-----------------|
| 28 | 62 | 5 | Vsi bez kostelu |
| 29 | 63 | 5 | Vsi bez kostelu |
| 30 | 69 | 7 | Vsi s kostely |
| 31 | 70 | 5 | Vsi bez kostelu |
| 32 | 71 | 5 | Vsi bez kostelu |
| 33 | 72 | 5 | Vsi bez kostelu |
| 34 | 152 | 5 | Vsi bez kostelu |
| 35 | 173 | 6 | Vsi se zamky |
| 36 | 194 | 5 | Vsi bez kostelu |
| 37 | 195 | 5 | Vsi bez kostelu |

4700 rows.

Počet řádků v tabulce MULLER_UNIQUE_ZAZNAMY je 4700, což přesně odpovídá počtu všech záznamů v tabulce SIDLA_MULLER (6278) bez počtu existujících sídel majících v tabulce SIDLA_MULLER dvojné uložení (812) a bez počtu všech záznamů zaniklých sídel (766). V datech tedy není chyba týkající se nesouhlasu počtu záznamů.

Následně byl použit výběrový příkaz, který ve dvou krocích sestavil tabulku POMER_UNIQUE. Nejprve jsem sestavil pomocnou tabulku POMER_UNIQUE_BASE a poté jsem ji spojil s tabulkou TYP_SIDEŁ. Toto je znázorněno v Příkazu 16.

Příkaz 16: Tvorba tabulky POMER_UNIQUE

```
CREATE TABLE pomer_unique_base AS(
  SELECT
    typ,
    COUNT (typ) AS pocet
  FROM
    muller_unique_zaznamy
  GROUP BY (typ)
  HAVING COUNT (typ) >= 1
);
CREATE TABLE pomer_unique AS(
  SELECT
    *
  FROM
    pomer_unique_base
  INNER JOIN
    typ_sidel
  USING (typ)
  ORDER BY (typ)
);
```

Tím byla vytvořena tabulka POMER_UNIQUE, jejíž popis je v Tabulce 19 a její ukázka na Obrázku 28.

Tabulka 19: Popis tabulky POMER_UNIQUE

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEŁ |
| pocet | bigint | počet nalezených záznamů stejného typu |
| vyznam | text | vysvětlení typu sídla |

Obrázek 28: Ukázka tabulky POMER_UNIQUE

| | typ integer | pocet bigint | vyznam text |
|----|----------------|-----------------|----------------------------------|
| 1 | 1 | 2 | Mestecka se zamky |
| 2 | 2 | 174 | Mestecka s trhy |
| 3 | 3 | 37 | Neopevnená mesta |
| 4 | 4 | 33 | Opevnená mesta |
| 5 | 5 | 3710 | Vsi bez kostelu |
| 6 | 6 | 139 | Vsi se zamky |
| 7 | 7 | 484 | Vsi s kostely |
| 8 | 8 | 33 | Vsi s kostely a zamky |
| 9 | 9 | 70 | Vsi s rytirskými sídly |
| 10 | 10 | 18 | Vsi s rytirskými sídly a kostely |

Výsledky uvedené v procentuálním zastoupení jsou popsány v Tabulce 20.

Tabulka 20: Poměr neredundantních sídel v tabulce POMER_UNIQUE

| Typ sídla | Procentuální zastoupení |
|---------------------------------------|-------------------------|
| 1 - Městečka se zámky | < 0,01 % |
| 2 - Městečka s trhy | 3,7 % |
| 3 - Neopevňená města | 0,7 % |
| 4 - Opevňená města | 0,7 % |
| 5 - Vsi bez kostelů | 78,9 % |
| 6 - Vsi se zámky | 3 % |
| 7 - Vsi s kostely | 10,3 % |
| 8 - Vsi s kostely a zámky | 0,7 % |
| 9 - Vsi s rytířskými sídly | 1,5 % |
| 10 - Vsi s rytířskými sídly a kostely | 0,4 % |

Případ 2: Poměr redundantních existujících sídel v tabulce SIDLA_MULLER

Dalším krokem bylo vytvoření tabulky obsahující duplicitní záznamy, ale s určitými podmínkami. Způsob tvorby s podmínkami je uveden v Příkazu 17.

Příkaz 17: Tvorba tabulky MULLER_REDUNDANTNI_ZAZNAMY

```
CREATE TABLE muller_redundantni_zaznamy AS (
SELECT DISTINCT ON (s.id)
    a.typ,
    s.id AS id_sidla
FROM   sidla_muller a,
       sidla_muller b,
       sidla s
WHERE  a.id_sidla = b.id_sidla AND
       a.id <> b.id AND
       s.id = a.id_sidla AND
       a.typ = b.typ AND
       a.id_sidla < 10000000
);
```

Příkaz 17 umožnil vytvořit tabulku, která obsahuje všechna duplicitně uložená sídla v tabulce SIDLA_MULLER, která v dnešní době existují, což značí hodnota `id_sidla` v tabulce MULLER_REDUNDANTNI_ZAZNAMY, která je menší než 10000000. Avšak tato sídla již nebyla uložena dvojně, ale každé jednotlivé mělo jen jeden záznam, protože podvojně záznamy obsahovaly stejné hodnoty ve sloupci `typ` (nenastane tudíž situace vyobrazená na Obrázku 26).

Tabulka MULLER_REDUNDANTNI_ZAZNAMY je popsána v Tabulce 21.

Tabulka 21: Popis tabulky MULLER_REDUNDANTNI_ZAZNAMY

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEL |
| id_sidla | integer | id sídel uvedených v tabulce SIDLA |

Část tabulky MULLER_REDUNDANTNI_ZAZNAMY je zobrazena na Obrázku 29.

Obrázek 29: Část vyobrazení tabulky MULLER_REDUNDANTNI_ZAZNAMY

| | typ integer | id_sidla bigint |
|-----|----------------|--------------------|
| 99 | 5 | 4684 |
| 100 | 7 | 4695 |
| 101 | 5 | 4696 |
| 102 | 5 | 4702 |
| 103 | 5 | 4707 |
| 104 | 5 | 4711 |
| 105 | 7 | 4712 |
| 106 | 5 | 4714 |
| 107 | 7 | 4716 |
| 108 | 5 | 4717 |
| 109 | 5 | 4718 |

327 rows.

Nyní vytvořím tabulku POMER_REDUNDANTNI opět ve dvou krocích jako v Případě 1. Způsob je popsán v Příkazu 18.

Příkaz 18: Tvorba tabulky POMER_REDUNDANTNI

```
CREATE TABLE pomer_redundantni_base AS(
  SELECT
    typ,
    COUNT (typ) AS pocet
  FROM
    muller_redundantni_zaznamy
    GROUP BY (typ)
    HAVING COUNT (typ) >= 1
);
CREATE TABLE pomer_redundantni AS(
  SELECT
    *
  FROM
    pomer_redundantni_base
  INNER JOIN
    typ_sidel
  USING (typ)
  ORDER BY (typ)
);
```

Ukázka tabulky POMER_REDUNDANTNI je na Obrázku 30 její popis v Tabulce 22. Co se týká sídel, která ve dvojně uložených záznamech měla rozdílné označení typu tohoto sídla, ta byla ze statistiky vynechána. Avšak toto by mohlo být podrobeno zkoumání, proč tato označení jsou rozdílná. Mohlo by to být například díky vývoji sídel mezi mapováním jednotlivých mapových listů nebo změněného označování a značení typů sídel. To však již nebude předmětem mé práce.

Tabulka 22: Popis tabulky POMER_REDUNDANTNI

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEL |
| pocet | bigint | počet nalezených záznamů stejného typu |
| vyznam | text | vysvětlení typu sídla |

Obrázek 30: Zobrazení tabulky POMER_REDUNDANTNI

| | typ integer | pocet bigint | vyznam text |
|---|----------------|-----------------|----------------------------------|
| 1 | 2 | 20 | Mestecka s trhy |
| 2 | 3 | 1 | Neopevnená mesta |
| 3 | 4 | 8 | Opevnená mesta |
| 4 | 5 | 270 | Vsi bez kostelu |
| 5 | 6 | 2 | Vsi se zamky |
| 6 | 7 | 21 | Vsi s kostely |
| 7 | 9 | 4 | Vsi s rytirskými sídly |
| 8 | 10 | 1 | Vsi s rytirskými sídly a kostely |

Poměr sídel v procentech je uveden v Tabulce 23.

Tabulka 23: Poměr všech sídel v tabulce POMER_REDUNDANTNI

| Typ sídla | Procentuální zastoupení |
|---------------------------------------|-------------------------|
| 2 - Městečka s trhy | 6,1 % |
| 3 - Neopevňená města | 0,3 % |
| 4 - Opevňená města | 2,4 % |
| 5 - Vsi bez kostelů | 82,6 % |
| 6 - Vsi se zámky | 0,6 % |
| 7 - Vsi s kostely | 6,4 % |
| 9 - Vsi s rytířskými sídly | 1,2 % |
| 10 - Vsi s rytířskými sídly a kostely | 0,3 % |

Případ 3: Sloučení Případu 1 a Případu 2

Nyní přistoupím k tomu, že zhotovím tabulku, znázorňující ucelený komplet dat a to tím způsobem, že spojím tabulku MULLER_UNIQUE_ZAZNAMY a MULLER_REDUNDANTNI_ZAZNAMY do jedné souhrnné tabulky nazvané MULLER_UNIQUE_I_REDUNDANTNI. Následně vytvořím tabulku POMER_VSECH, díky níž bude možno získat statistiku zahrnující maximální možný počet dnes existujících sídel při dodržení podmínek kladených na výběr sídel již v Případě 1 a Případě 2.

Sloučení proběhlo dotazem uvedeným v Příkazu 19.

Příkaz 19: Vytvoření tabulky MULLER_UNIQUE_I_REDUNDANTNI

```
CREATE TABLE muller_unique_i_redundantni AS (  
  SELECT  
    typ,  
    id_sidla  
  FROM  
    muller_unique_zaznamy  
  UNION  
  SELECT  
    typ,  
    id_sidla  
  FROM  
    muller_redundantni_zaznamy  
);
```

Ukázka tabulky MULLER_UNIQUE_I_REDUNDANTNI je uvedena na Obrázku 31. Její struktura je popsána v Tabulce 21.

Obrázek 31: Ukázka tabulky MULLER_UNIQUE_I_REDUNDANTNI

| | typ integer | id_sidla bigint |
|-----------|-----------------------|---------------------------|
| 1 | 1 | 10064 |
| 2 | 1 | 16879 |
| 3 | 2 | 394 |
| 4 | 2 | 536 |
| 5 | 2 | 680 |
| 6 | 2 | 2816 |
| 7 | 2 | 4024 |
| 8 | 2 | 4081 |
| 9 | 2 | 4584 |
| 10 | 2 | 4720 |
| 11 | 2 | 5430 |

5027 rows.

Počet uložených záznamů (5027) odpovídá součtu záznamů v tabulce MULLER_UNIQUE_ZAZNAMY (4700) a MULLER_REDUNDANTNI_ZAZNAMY (327).

Nyní vytvořím tabulku POMER_VSECH souhrně ukazující zastoupení sídel v tabulce MULLER_UNIQUE_I_REDUNDANTNI dle jejich typů. Způsob vytvoření je uveden v Příkazu 20.

Příkaz 20: Tvorba tabulky POMER_VSECH

```
CREATE TABLE pomer_vsech_base AS(  
  SELECT  
    typ,  
    COUNT (typ) AS pocet  
  FROM  
    muller_unique_i_redundantni  
  GROUP BY (typ)  
  HAVING COUNT (typ) >= 1  
);  
CREATE TABLE pomer_vsech AS(  
  SELECT  
    *  
  FROM  
    pomer_vsech_base  
  INNER JOIN  
    typ_sidel  
  USING (typ)
```


ORDER BY (typ)
);

Popis tabulky POMER_VSECH je zobrazen v Tabulce 24 a její část je vyobrazena na Obrázku 32. Samotný procentuální přehled zastoupení typů sídel je uveden v Tabulce 25.

Tabulka 24: Popisek tabulky POMER_VSECH

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEL |
| pocet | bigint | počet nalezených záznamů stejného typu |
| vznam | text | vysvětlení typu sídla |

Obrázek 32: Přehled tabulky POMER_VSECH

| | typ integer | pocet bigint | vznam text |
|-----------|----------------|-----------------|----------------------------------|
| 1 | 1 | 2 | Mestecka se zamky |
| 2 | 2 | 194 | Mestecka s trhy |
| 3 | 3 | 38 | Neopevnená mesta |
| 4 | 4 | 41 | Opevnená mesta |
| 5 | 5 | 3980 | Vsi bez kostelu |
| 6 | 6 | 141 | Vsi se zamky |
| 7 | 7 | 505 | Vsi s kostely |
| 8 | 8 | 33 | Vsi s kostely a zamky |
| 9 | 9 | 74 | Vsi s rytirskými sídly |
| 10 | 10 | 19 | Vsi s rytirskými sídly a kostely |

Tabulka 25: Poměr všech sídel v tabulce POMER_VSECH

| Typ sídla | Procentuální zastoupení |
|---------------------------------------|-------------------------|
| 1 - Městečka se zámky | < 0,01 % |
| 2 - Městečka s trhy | 3,9 % |
| 3 - Neopevňená města | 0,8 % |
| 4 - Opevňená města | 0,8 % |
| 5 - Vsi bez kostelů | 79,1 % |
| 6 - Vsi se zámky | 2,8 % |
| 7 - Vsi s kostely | 10 % |
| 8 - Vsi s kostely a zámky | 0,7 % |
| 9 - Vsi s rytířskými sídly | 1,5 % |
| 10 - Vsi s rytířskými sídly a kostely | 3,6 % |

Případ 4: Poměr již zaniklých sídel v tabulce SIDLA_MULLER

Nyní sestavím tabulku MULLER_ZANIKLA_SIDLA, obsahující pouze dnes již neexistující sídla. Způsob vytvoření popsán v Příkazu 21.

Příkaz 21: Vytvoření tabulky MULLER_ZANIKLA_SIDLA

```
CREATE TABLE muller_zanikla_sidla AS (  
  SELECT DISTINCT ON (v.id_sidla)  
    v.id_sidla,  
    v.typ  
  FROM  
    vysoka_id v  
  WHERE  
    v.id_sidla IN  
    (  
      SELECT  
        id_sidla  
      FROM  
        vysoka_id  
      GROUP BY (id_sidla)  
      HAVING COUNT (id_sidla) = 1  
    )  
  
  UNION  
  SELECT DISTINCT ON (v.id_sidla)  
    v.id_sidla,  
    v.typ  
  FROM  
    vysoka_id v,  
    vysoka_id b  
  WHERE  
    v.id_sidla IN  
    (  
      SELECT  
        id_sidla  
      FROM  
        vysoka_id  
      GROUP BY (id_sidla)  
      HAVING COUNT (id_sidla) = 2  
    )  
  AND v.typ = b.typ  
);
```

Popis tabulky MULLER_ZANIKLA_SIDLA je uveden v Tabulce 26 a zobrazen na Obrázku 33. Tabulka VYSOKA_ID, zmíněná v SQL příkazu k vytvoření tabulky MULLER_ZANIKLA_SIDLA je jen jednoduchá tabulka, ve které jsou vybraná všechna osmimístná id z tabulky SIDLA, která značí nynější neexistenci těchto sídel.

Z údajů vyobrazených na Obrázku 33 lze pomocí rozkódování typu sídla dle tabulky TYP_SIDLA určit procentuální zastoupení typu zaniklých sídel. Celkem jednotlivých zaniklých sídel je v databázi uloženo 759 z celkového počtu 766. Sedm sídel má dvojí záznam a všechny dvojice mají stejný typ, nemusel jsem tedy z pohledu dvojznačnosti typu sídla žádné vyloučit, pouze jsem mohl vzít vždy jen jeden záznam z každé dvojice.

Tabulka 26: Popis tabulky MULLER_ZANIKLA_SIDLA

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEŁ |
| pocet | bigint | počet nalezených záznamů stejného typu |

Obrázek 33: Ukázka tabulky MULLER_ZANIKLA_SIDLA

| | id_sidla integer | typ integer |
|-----------|----------------------------|-----------------------|
| 49 | 10000050 | 5 |
| 50 | 10000051 | 2 |
| 51 | 10000052 | 2 |
| 52 | 10000053 | 6 |
| 53 | 10000054 | 7 |
| 54 | 10000055 | 7 |
| 55 | 10000056 | 5 |
| 56 | 10000057 | 5 |
| 57 | 10000058 | 5 |
| 58 | 10000059 | 5 |

759 rows.

Z dat uložených v tabulce MULLER_ZANIKLA_SIDLA sestavím tabulku POMER_ZANIKLA tak, jak je uvedeno v Příkazu 22.

Příkaz 22: Tvorba tabulky POMER_ZANIKLA

```
CREATE TABLE pomer_zanikla_base AS(
SELECT
    typ,
    COUNT (typ) AS pocet
FROM
    zanikla
    GROUP BY (typ)
    HAVING COUNT (typ) >= 1
);
CREATE TABLE pomer_zanikla AS(
SELECT
    *
FROM
    pomer_zanikla_base
INNER JOIN
    typ_sidel
USING (typ)
ORDER BY (typ)
);
```

Popis tabulky POMER_ZANIKLA naleznete v Tabulce 27 a její část je vyobrazena na Obrázku 34. Zobrazení přehledu zastoupení sídel je uvedeno v Tabulce 28.

Tabulka 27: Popis tabulky POMER_ZANIKLA

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| typ | integer | typ sídla korespondující s popisem uloženým v tabulce TYP_SIDEL |
| pocet | bigint | počet nalezených záznamů stejného typu |
| vyznam | text | vysvětlení typu sídla |

Obrázek 34: Přehled tabulky POMER_ZANIKLA

| | typ integer | pocet bigint | vyznam text |
|---|-------------|--------------|------------------------|
| 1 | 2 | 12 | Mestecka s trhy |
| 2 | 3 | 4 | Neopevnena mesta |
| 3 | 4 | 3 | Opevnena mesta |
| 4 | 5 | 674 | Vsi bez kostelu |
| 5 | 6 | 16 | Vsi se zamky |
| 6 | 7 | 45 | Vsi s kostely |
| 7 | 8 | 2 | Vsi s kostely a zamky |
| 8 | 9 | 3 | Vsi s rytirskymi sidly |

Tabulka 28: Poměr již zaniklých sídel v tabulce SIDLA_MULLER

| Typ sídla | Procentuální zastoupení |
|----------------------------|-------------------------|
| 2 - Městečka s trhy | 1,6 % |
| 3 - Neopevněná města | 0,5 % |
| 4 - Opevněná města | 0,4 % |
| 5 - Vsi bez kostelů | 88,8 % |
| 6 - Vsi se zámky | 2,1 % |
| 7 - Vsi s kostely | 5,9 % |
| 8 - Vsi s kostely a zámky | 0,3 % |
| 9 - Vsi s rytířskými sídly | 0,4 |

Tabulky 20, 23, 25 a 28 můžeme vzájemně porovnat (viz. Tabulka 29) a vyřknout celkový závěr k této statistice.

Tabulka 29: Celkový přehled zastoupení typů sídel

| Typ sídla | Případ 1 | Případ 2 | Případ 3 | Případ 4 |
|---------------------------------------|----------|----------|----------|----------|
| 1 - Městečka se zámky | < 0,01 % | | < 0,01 % | |
| 2 - Městečka s trhy | 3,7 % | 6,1 % | 3,9 % | 1,6 % |
| 3 - Neopevněná města | 0,7 % | 0,3 % | 0,8 % | 0,5 % |
| 4 - Opevněná města | 0,7 % | 2,4 % | 0,8 % | 0,4 % |
| 5 - Vsi bez kostelů | 78,9 % | 82,6 % | 79,1 % | 88,8 % |
| 6 - Vsi se zámky | 3 % | 0,6 % | 2,8 % | 2,1 % |
| 7 - Vsi s kostely | 10,3 % | 6,4 % | 10 % | 5,9 % |
| 8 - Vsi s kostely a zámky | 0,7 % | | 0,7 % | 0,3 % |
| 9 - Vsi s rytířskými sídly | 1,5 % | 1,2 % | 1,5 % | 0,4 % |
| 10 - Vsi s rytířskými sídly a kostely | 0,4 % | 0,3 % | 3,6 % | |

Z Tabulky 29 je jednoznačně určitelných několik faktů:

1. Městečka se zámky se vyskytují pouze v souboru unikátně uložených sídel (Případ 1), tím pádem se na mapových listech vyskytují v části, která se nepřekrývá s jinými listy.
2. V rámci zaniklých sídel dominuje počet Vsí bez kostelů, který ostatně dominuje ve všech případech, v Případě 4 však největší procentuální mírou.
3. V rámci zaniklých sídel, se neobjevilo žádné sídlo označené jako Ves s rytířskými sídly a kostely, což značí to, že tato sídla fakticky existují dodnes, i když ne se stejným označením.
4. Některá odlišná čísla v procentech v Případě 2 oproti Případu 1 jsou způsobena rozsahem souboru, který byl pro Případ 2 14-krát menší než v Případě 1, takže by bylo možno z procentuálních hodnot v Případě 2 vyřknout závěr, že se mezi nimi vyskytoval větší počet opevněných měst než v Případě 1, ale co do absolutních čísel to pravda není.

Z celkového pohledu na Tabulku 29 plyne, že v době vzniku rukopisných Müllerových map bylo rozložení typu sídel dáno jak sociálními tak i politickými aspekty, ať už jde o movitost jedinců, šlechty, církve, o věhlas sídla, tak i o význam umístění sídel. Sídel s významnými stavbami (kostely, zámky, rytířská sídla) je tudíž mnohonásobně méně než "obyčejných vesnických".

3.2.4 Vývoj typů sídel mezi počátkem 18. století a současností

Další analýzou dat bude zjištění vývoje typů sídel v rámci současného a tehdejšího označování.

Nynější označení sídel samozřejmě nekoresponduje s tím, které bylo používáno na počátku 18. století. Účelem je tudíž zjistit, zda je nějaká spojitost v tom, které typy sídel se měnily na ty, jejichž označení se používá dnes. Předem bylo tedy nutno poupravit strukturu databáze, která tomuto účelu v podobě, v jaké byla do tohoto momentu používána, nevyhovovala.

Proto byl do tabulky SIDLA přidán další sloupec `csu_typ`, který nese informaci o označení typu sídla, poskytnutého ČSÚ, vztahujícím se k roku 2008. Jedná se o databázové soubory zmíněné v Kapitole 2.1, nyní uvedu přehled jejich struktury dle [2] v Tabulce 30.

Tabulka 30: Přehled užitých databázových souborů

| Databázový soubor | Vybrané atributy |
|-------------------|---|
| OBCE.dbf | Název obce |
| | Souřadnice definičního bodu obce (Y, X) |
| | Kód obce |
| KU.dbf | Název k.ú. |
| | Souřadnice definičního bodu k.ú. (Y, X) |
| | Kód obce |
| | Kód k.ú. |
| MCAST.dbf | Název MO/MČ |
| | Souřadnice definičního bodu MO/MČ (Y, X) |
| | Kód obce |
| | Kód MO/MČ |
| COBE.dbf | Název části obce evidenční |
| | Souřadnice definičního bodu části obce evidenční (Y, X) |
| | Kód části obce evidenční |
| COB.dbf | Název části obce |
| | Souřadnice definičního bodu části obce (Y, X) |
| | Kód obce |
| | Kód části obce |
| | Kód části obce evidenční |
| ZSJ.dbf | Název ZSJ |
| | Souřadnice definičního bodu ZSJ (Y, X) |
| | Kód obce |
| | Kód k.ú. |
| | Kód MO/MČ |
| | Kód ZSJ |

Dle Tabulky 30, kde v každém "xxx.dbf" souboru byla uvedena sídla o stejném typu, jsem aktualizoval i tabulku SIDLA v databázi. Prvotně jsem sloupec `csu_typ` naplnil implicitní hodnotou řetězce 'xxx', značící 'neurčeno', a postupně jsem tabulku sídla aktualizoval Příkazem 23.

Příkaz 23: Finální aktualizace tabulky SIDLA

```
UPDATE
    sidla
SET
    csu_typ = 'zsj'
FROM
    zsj
WHERE
    zsj.sx = sidla.x_puv AND
    zsj.sy = sidla.y_puv AND
    sidla.csu_typ = 'xxx'
;
```

Přičemž tato ukázka příkazu aktualizující sloupec `csu_typ` je poslední dílčí částí aktualizace tabulky SIDLA. Ta probíhala přesně v pořadí souborů, jak je uvedeno v Tabulce 30 (shora dolů), jen se příkaz měnil na příslušných místech, kde byl zanesen název vstupního souboru, respektive hodnota aktualizovaná do sloupce `csu_typ` tabulky SIDLA. Takto bylo aktualizováno všech 22215 sídel získaných od ČSÚ. Ukázku konečné podoby tabulky SIDLA naleznete na Obrázku 35, její strukturu na Obrázku 36 a popis v Tabulce 32.

Obrázek 35: Finální podoba tabulky SIDLA pro tuto práci

| | id [PK] bigint | nazev_cz text | x_puv double precis | y_puv double precis | x_nova double precis | y_nova double precis | zmeneno integer | csu_typ text |
|------|-------------------|------------------|---------------------------|---------------------------|----------------------------|----------------------------|--------------------|-----------------|
| 5453 | 5453 | Libčice | 1078585 | 755872 | 1078585 | 755872 | 1 | ku |
| 5454 | 5454 | Starý Knín | 1074670 | 757189 | 1074670 | 757189 | 1 | ku |
| 5455 | 5455 | Sudovice | 1075047 | 755258 | 1075101,85 | 755315,85 | 1 | ku |
| 5456 | 5456 | Oseč | 1080657 | 781593 | 1080614,54 | 781735,49 | 1 | ku |
| 5457 | 5457 | Lhotka_u_Dobří | 1076843 | 766802 | 1076828,64 | 766794,68 | 1 | ku |
| 5458 | 5458 | Velběhy | 1084585 | 746050 | 1084617,44 | 746137,98 | 1 | ku |
| 5459 | 5459 | Bedřichov | 1130215 | 605263 | 1130081 | 605164 | 1 | obce |
| 5460 | 5460 | Běleč | 1132571 | 611519 | 1132524 | 611524 | 1 | obce |
| 5461 | 5461 | Benešov | 1127254 | 582697 | 1127308 | 582863 | 1 | obce |
| 5462 | 5462 | Borotín | 1118541 | 588981 | 1118432 | 589049 | 1 | obce |
| 5463 | 5463 | Bořitov | 1135201 | 596626 | 1135063 | 596654 | 1 | obce |
| 5464 | 5464 | Boskovice | 1128828 | 590923 | 1128719 | 591013 | 1 | obce |
| 5465 | 5465 | Brťov-Jeneč | 1135828 | 602168 | 1135764 | 602254 | 1 | obce |

U sídel, která mají ve sloupci id tabulky SIDLA osmimístnou hodnotu, tzn. nyní již neexistují, nebyla poskytnuta data z ČSÚ. Proto u těchto sídel je ve sloupci csu_typ uvedena implicitní hodnota značící neurčení tohoto sídla z pohledu dnešního označení typu sídel.

Z doplnění dalšího sloupce do tabulky SIDLA a jeho aktualizace vyplývá, že se změnila struktura databáze. Přibyla další databázová tabulka TYP_CSU, jejíž popis je uveden v Tabulce 31 a do tabulky SIDLA se přiřadil další sloupec.

Tabulka 31: Popis tabulky TYP_CSU

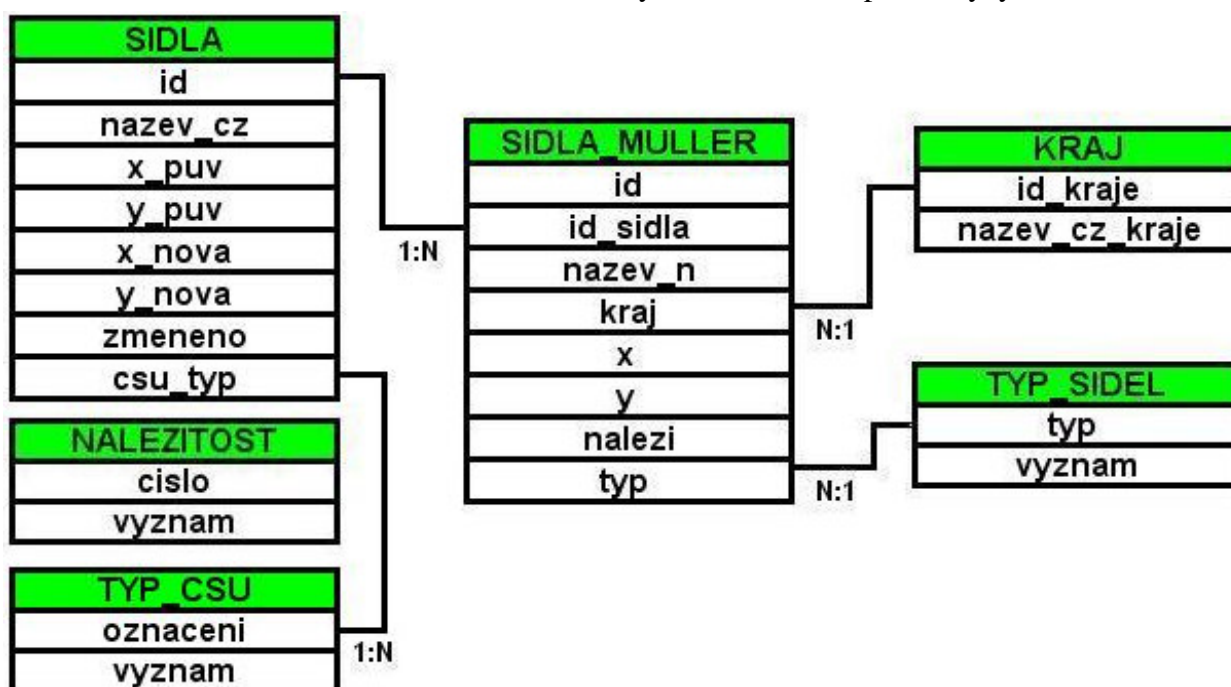
| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| oznaceni | text | označení typu sídla přiřazený z dat ČSÚ, cizí klíč k tabulce SIDLA, k sloupci csu_typ |
| vyznam | text | klíč k určení významu sídla |

Ukázka tabulky TYP_CSU je uvedena na Obrázku 36.

Obrázek 36: Přehled tabulky TYP_CSU

| | oznaceni [PK] text | vyznam text |
|---|-----------------------|---|
| 1 | cob | casti obci |
| 2 | cobe | casti obci evidencnich |
| 3 | ku | katastralni uzemi |
| 4 | mcast | mestske casti a obvody |
| 5 | obce | definicni body obci a vojenskych ujezdu |
| 6 | xxx | neurceno |
| 7 | zsj | zakladni sidelni jednotky |
| * | | |

Obrázek 37: Finální struktura výchozích tabulek pro analýzy



Tabulka 32: Konečná podoba tabulky SIDLA

| Název atributu | Datový typ PostgreSQL | Popis |
|----------------|-----------------------|---|
| id | bigint | jinak také označen jako primární klíč, slouží pro přístup k datům (entitám tabulky) |
| nazev_cz | text | název sídla v českém jazyce |
| x_puv | double precision | původní X-ová souřadnice sídla poskytnutá ČSÚ |
| y_puv | double precision | původní Y-ová souřadnice sídla poskytnutá ČSÚ |
| x_nova | double precision | aktualizovaná X-ová souřadnice určená skrze ZM10 [2], [3] |
| y_nova | double precision | aktualizovaná Y-ová souřadnice určená skrze ZM10 [2], [3] |
| zmeneno | integer | popis v Tabulce 5 |
| csu_typ | text | označení typu sídla v rámci dat poskytnutých ČSÚ, implicitní hodnota 'xxx' |

Nyní byla databáze dostatečně funkční a transparentní na to, aby se provedla analýza, kterou se tato kapitola zabývá. Nejdříve jsem tedy vybral ta odpovídající si sídla mezi tabulkami SIDLA a SIDLA_MULLER, kde se vyskytovalo ve sloupci csu_typ označení 'obce'. K tomu posloužil Příkaz 24.

Příkaz 24: Vzor pro tvorbu tabulky SROVNANI_OBCE a jí podobných

```
CREATE TABLE srovnani_obce_base AS (
    SELECT
        sidla.csu_typ ,
        muller_unique_i_redundantni.typ
    FROM
        sidla,
        muller_unique_i_redundantni
    WHERE
        csu_typ = 'obce' AND
        sidla.id = muller_unique_i_redundantni.id_sidla)
;
CREATE TABLE srovnani_obce AS (
    SELECT
        typ,
        COUNT (typ) AS pocet
    FROM
        srovnani_obce_base
    GROUP BY (typ)
    HAVING COUNT (typ) >= 1
    ORDER BY (typ)
);
```

Tento příkaz jsem nechal vykonat 6-krát, pokaždé jsem změnil název tabulky (SROVNANI_OBCE_BASE,SROVNANI_KU_BASE, atd.) a dle toho jsem měnil i označení v klauzuli

```
WHERE
    csu_typ = 'obce'...
```

Poté jsem použil Příkaz 25 ke srovnání počtu a typů sídel, mající v dnešní době jedno ze šesti výše uvedených označení.

Příkaz 25: Vytvoření výběrů z tabulek SROVNANI_XXX

```
SELECT
*
FROM
  srovnani_obce
INNER JOIN
  typ_sidel
  USING (typ)
;
```

Opětovným zaměňováním názvu tabulky, ze které jsem vybíral, uvedené v Příkazu 24, jsem získal 6 tabulek, které jsou uvedeny jako Obrázky 38 až 43.

Obrázek 38: Přehled tabulky SROVNANI_OBCE

| | typ integer | pocet bigint | vyznam text |
|----|----------------|-----------------|----------------------------------|
| 1 | 1 | 2 | Mestecka se zamky |
| 2 | 2 | 181 | Mestecka s trhy |
| 3 | 3 | 38 | Neopevnena mesta |
| 4 | 4 | 38 | Opevnena mesta |
| 5 | 5 | 1311 | Vsi bez kostelu |
| 6 | 6 | 83 | Vsi se zamky |
| 7 | 7 | 351 | Vsi s kostely |
| 8 | 8 | 28 | Vsi s kostely a zamky |
| 9 | 9 | 46 | Vsi s rytirskymi sidly |
| 10 | 10 | 16 | Vsi s rytirskymi sidly a kostely |

Obrázek 39: Přehled tabulky SROVNANI_KU

| | typ integer | pocet bigint | vyznam text |
|---|----------------|-----------------|----------------------------------|
| 1 | 2 | 10 | Mestecka s trhy |
| 2 | 4 | 3 | Opevnena mesta |
| 3 | 5 | 1945 | Vsi bez kostelu |
| 4 | 6 | 44 | Vsi se zamky |
| 5 | 7 | 128 | Vsi s kostely |
| 6 | 8 | 5 | Vsi s kostely a zamky |
| 7 | 9 | 19 | Vsi s rytirskymi sidly |
| 8 | 10 | 2 | Vsi s rytirskymi sidly a kostely |

Obrázek 40: Přehled tabulky SROVNANI_MCAST

| | typ integer | pocet bigint | vyznam text |
|--|----------------|-----------------|----------------|
| | | | |

Obrázek 41: Přehled tabulky SROVNANI_COBE

| | typ integer | pocet bigint | vznam text |
|---|----------------|-----------------|----------------------------------|
| 1 | 2 | 2 | Mestecka s trhy |
| 2 | 5 | 632 | Vsi bez kostelu |
| 3 | 6 | 13 | Vsi se zamky |
| 4 | 7 | 22 | Vsi s kostely |
| 5 | 9 | 6 | Vsi s rytirskymi sidly |
| 6 | 10 | 1 | Vsi s rytirskymi sidly a kostely |

Obrázek 42: Přehled tabulky SROVNANI_COB

| | typ integer | pocet bigint | vznam text |
|---|----------------|-----------------|------------------------|
| 1 | 5 | 1 | Vsi bez kostelu |
| 2 | 9 | 1 | Vsi s rytirskymi sidly |

Obrázek 43: Přehled tabulky SROVNANI_ZSJ

| | typ integer | pocet bigint | vznam text |
|---|----------------|-----------------|------------------------|
| 1 | 2 | 1 | Mestecka s trhy |
| 2 | 5 | 91 | Vsi bez kostelu |
| 3 | 6 | 1 | Vsi se zamky |
| 4 | 7 | 4 | Vsi s kostely |
| 5 | 9 | 2 | Vsi s rytirskymi sidly |

Jednoduchým porovnáním údajů zjistím to, že majoritní část bývalých sídel označených jako 5 - Vsi bez kostelů jsou v dnešní době označeny jako definiční bod katastrálního území, následovány těmi, které jsou zařazené mezi obce. Dalším zjevným faktem je absence sídel, které by byly nyní označeny jako městské části a jediná dvě sídla v databázi označená jako 1 - Městečka se zámky jsou nyní obcemi. Částmi obcí jsou pouhá dvě sídla.

Součty sídel náležejících k jedné ze 6 skupin dnešního dělení sídel jsou chronologicky uvedeny v Tabulce 33.

Tabulka 33: Součet sídel dle kategorií ČSÚ

| Označení | Počet sídel v tabulce SIDLA_MULLER mající označení dle ČSÚ |
|----------|--|
| KU | 2156 |
| OBCE | 2094 |
| COBE | 676 |
| ZSJ | 99 |
| COB | 2 |
| MCAST | 0 |

Vzhledem k počtu záznamů jednotlivých kategorií typů sídel v tabulce SIDLA, který je uveden v Tabulce reflektující aktualizaci sloupce `csu_typ` a dle [2] , nejsou výsledky velkým překvapením. Avšak vzhledem k poměru zastoupení typů sídel v tabulce SIDLA jsem očekával větší počet odpovídajících záznamů v tabulce SIDLA_MULLER pro základní sídelní jednotky. Ale vzhledem k definici ZSJ [8], uvozující jejich využití převážně ke statistickým účelům je zjevné, že toto rozdělení nebylo na počátku 18. století rozšířeno, respektive jen 99 záznamů v tabulce SIDLA_MULLER odpovídá této kategorii. Součet všech sídel v tabulce SIDLA_MULLER mající označení dle ČSÚ díky relaci s tabulkou SIDLA je 5027, což odpovídá počtu záznamů v tabulce MULLER_UNIQUE_I_REDUNDANTNI. Žádné sídlo tudíž nebylo z této tabulky vynecháno.

Závěr

První část této práce se zabývala tvorbou a údržbou databáze. Určení jejího účelu vycházelo z praktického hlediska dalšího jejího využití pro danou práci, avšak její struktura byla navržena tak, aby databáze mohla být použitelná i v budoucnu. To z ní tvoří vhodný a použitelný podklad pro další analýzy, ať už jde například o její rozšíření daty z jiného časového období mapování nebo pro lepší retrospektivní analýzy získaných údajů, tak i pro jiné oblasti zájmu.

Polohově určené analýzy, čili v mém případě souřadnicové analýzy, potvrdily domněnky o velikosti odchylek polohy sídel v rámci souřadnicové soustavy S-JTSK. Dalo by se říci, že ovlivňujícími faktory byly dvě zásadní věci: způsob mapování na počátku 18. století a vliv transformace souřadnic, i když vliv transformace je, předpokládám, výrazně menší. Toto hodnocení však vychází jen z malého vzorku Müllerových rukopisných map (viz. Příloha A). Bude moci být potvrzeno až v případě, kdy databáze bude plošná a bezešvá.

Atributové statistiky, zabývající se v jednom směru označením typů sídel v minulosti a v přítomnosti a ve druhém směru jednoduchým vývojem názvů zobrazených sídel, uvádějí přehled o těchto jevech a samotné statistiky lze v budoucnu rozšířit o další způsoby využití, respektive o rozšíření a o určení dalších výsledků z nich vyplývajících. Mohlo by se jednat například o tvorbu porovnávacího slovníku mezi českými a německými názvy sídel a jeho použití.

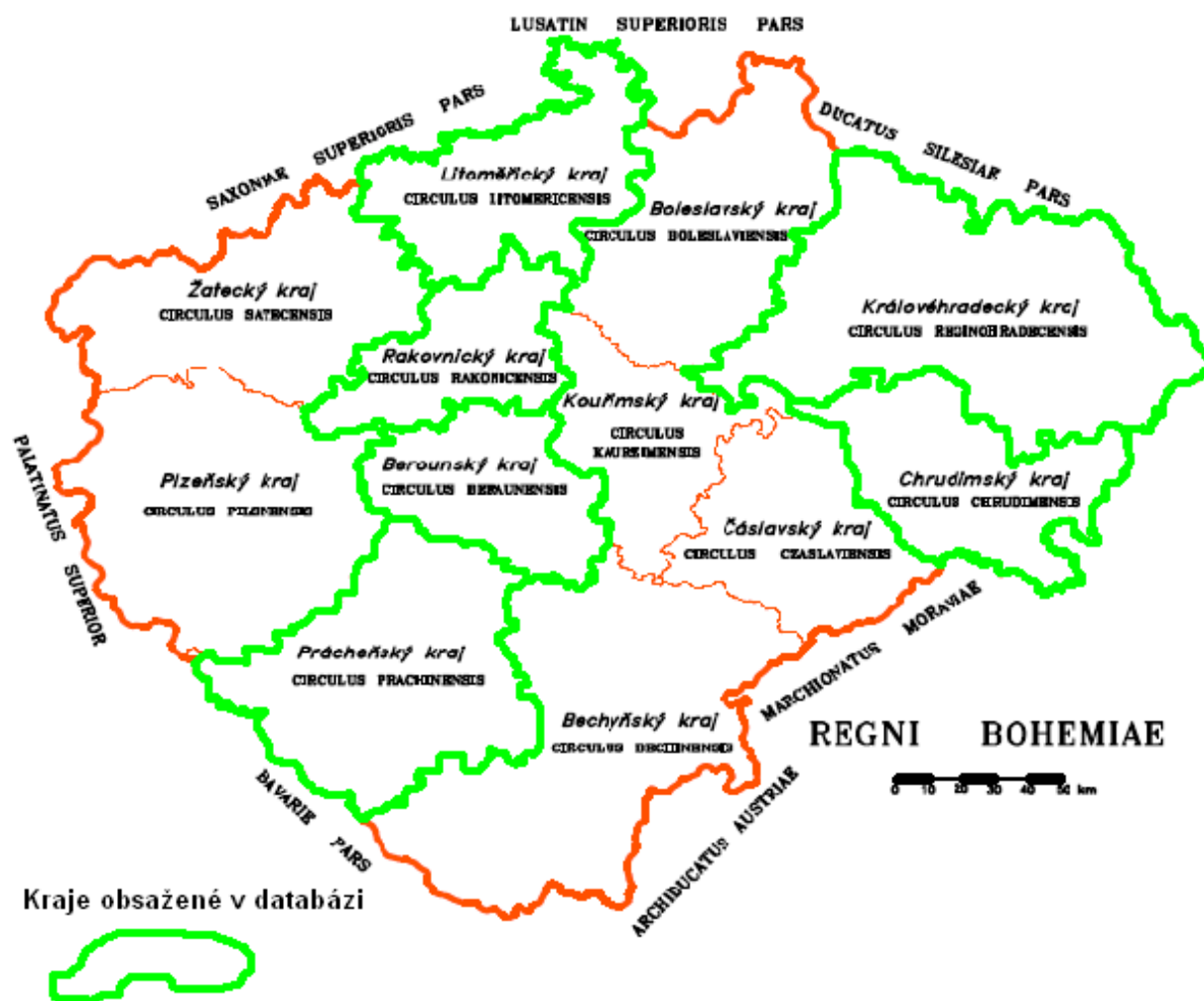
Jedná se tedy o základ a částečné využití potenciálu jak samotné vybudované databáze, tak i provedených analýz, což poskytuje prostor k dalšímu rozvoji. Konkrétně předpokládám doplnění databáze o data z dalších krajů z rukopisných Müllerových map Čech a tudíž vybudování celistvé databáze, kde by provedené analýzy měly komplexnější vypovídající efekt nejen o přesnosti použitých mapových děl. Dalším krokem by mohlo být rozšíření zdrojů, ze kterých by se data čerpala. Jednalo by se buď o rozšíření databáze o rukopisné mapy dalších území, nejen Čech, nebo o použití mapových zdrojů z jiných období, což by umožnilo provedení analýz jevů měnících se na daném území v čase. Celkový počet vytvořených a použitých tabulek je 30 (jejich přehled je zobrazen v Příloze C). Při následné aktualizaci databáze sídel dalšími daty z lokalizace Müllerových rukopisných map Čech a použitím stejných analýz, jako jsem použil v této práci já, se tento počet nezmění. Avšak díky předpokládanému rozšíření databáze o další mapové zdroje mimo Müllerových rukopisných map se o konečné číslo nejedná.

Formát vstupních dat vytvořený v rámci prací studentů zabývajících se lokalizací Müllerových rukopisných map, který byl použitý i jako zdroj do databáze sídel a byl zhotoven pro co nejefektivnější práci při samotné lokalizaci sídel. Vzhledem k tomu, že databáze sídel vychází z těchto dat, nikoli naopak, myslím si, že setrvání v používaném způsobu formátování dat, pro vkládání informací do tabulky SIDLA_MULLER, je použitelné i nadále.

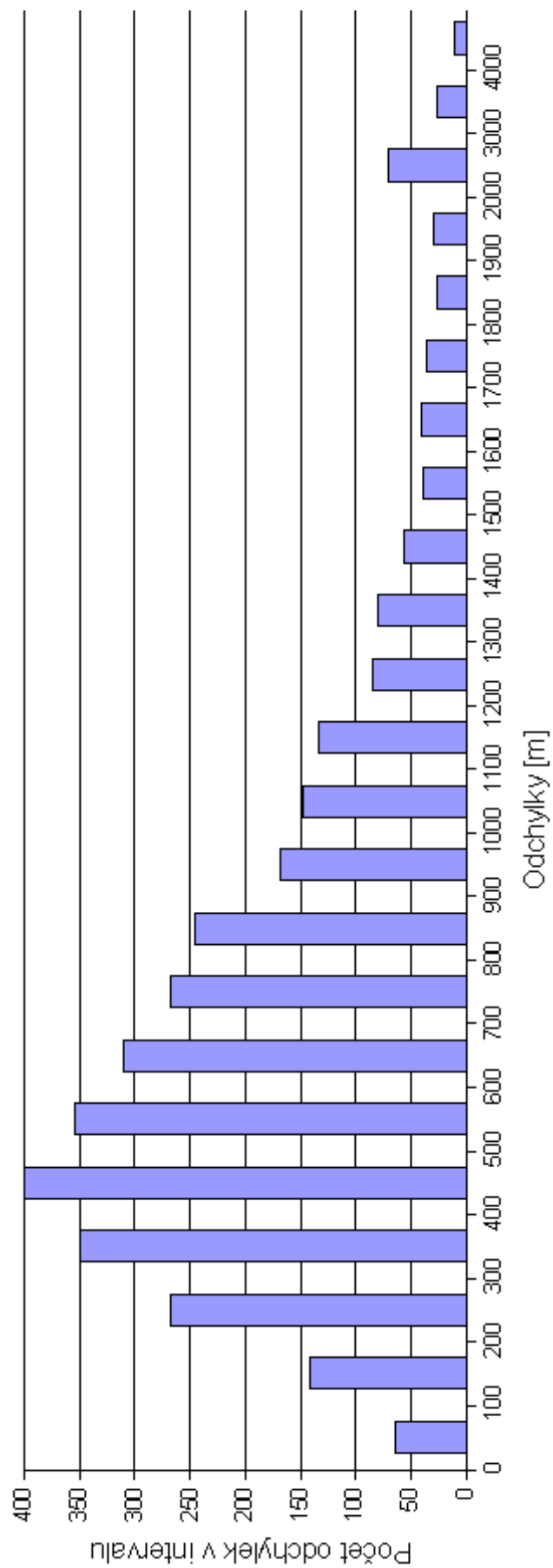
Při vypracování této bakalářské práce mi právě práce s databází poskytla prostor k proniknutí do problematiky tvorby, správy a údržby databází, jakožto nové oblasti poznání a vzdělání. Zároveň k tomu mohu i připočítat rozšíření znalostí o dotazovacím jazyku PostgreSQL a o způsobu jeho použití při práci s relačními databázemi.

Seznam příloh

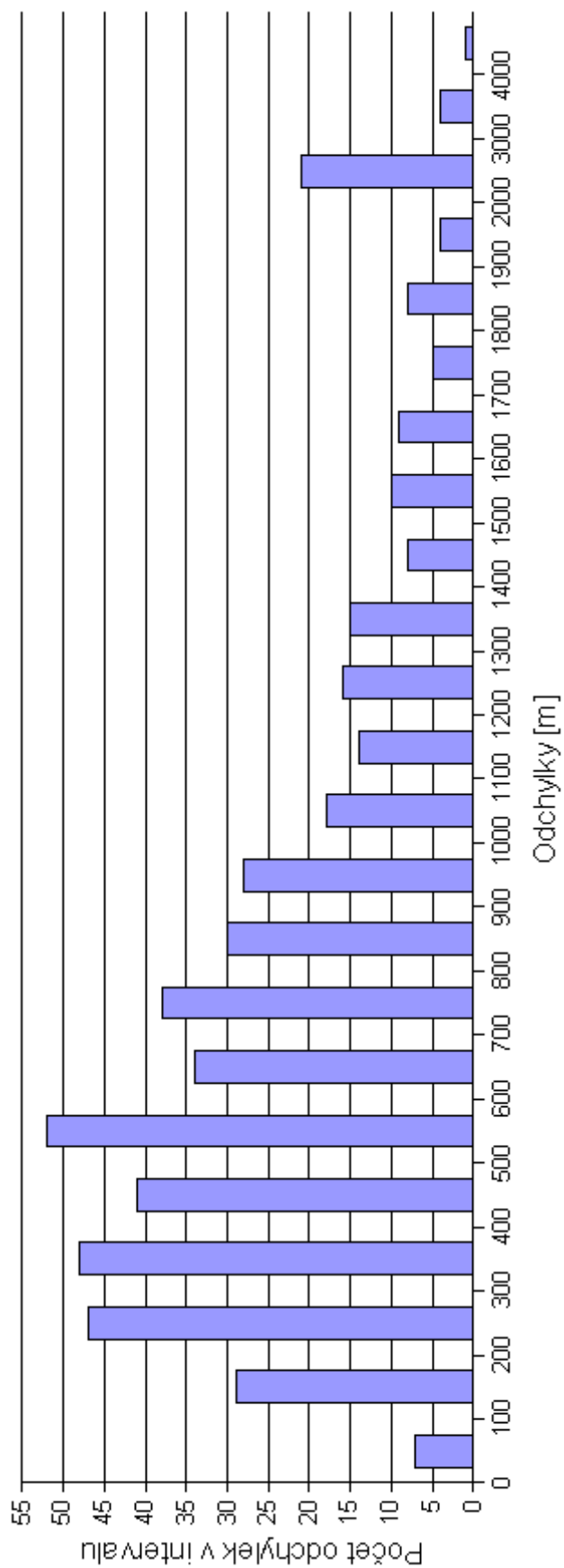
Příloha A: Přehled krajů Čech znázorněných na Müllerových rukopisných mapách [9]



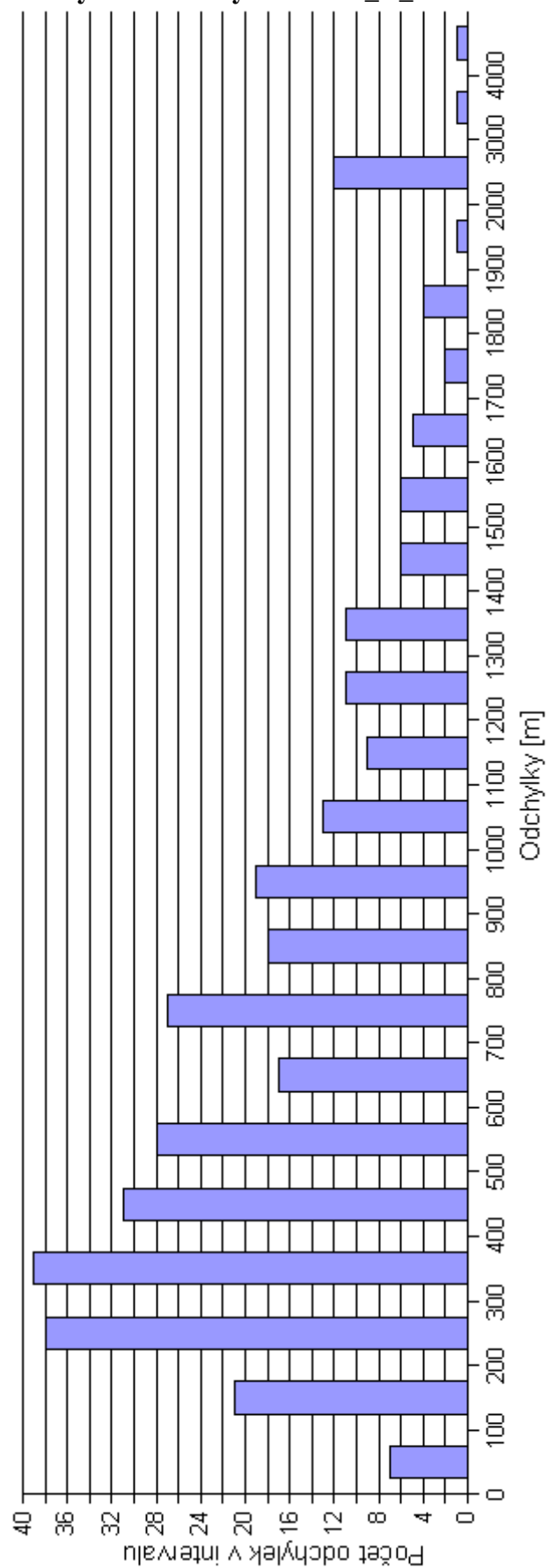
Příloha 1: Výběr odchylek tabulky ROZDIL_SIDLA_MULLER_UNIKATNI



Příloha 2 : Výběr odchylek tabulky ROZDIL_SIDLA_MULLER_DUPLICITNI



Příloha 3 : Výběr odchylek z tabulky ROZDIL_V_MULLEROVI_UNIKATNE



Příloha B: Obsah příloženého CD

Hájek_BP_2009

Databázové_soubory_ČSÚ

COB.DBF,
COBE.DBF,
KU.DBF,
MCAST.DBF,
OBCE.DBF,
ZSJ.DBF
a jejich SQL podoby

Kraj_BER_PRA

!_BER_vsechna_sidla_tps.stx - data pro aktualizaci tabulky SIDLA_MULLER
!_PRA_vsechna_sidla_tps.stx - data pro aktualizaci tabulky SIDLA_MULLER
BER_sidla_muller.sql - správný formát dat pro aktualizaci
PRA_sidla_muller.sql - správný formát dat pro aktualizaci
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

BER_PRA_typy_sidel

Soubory obsahující sídla dle jejich typu a jejich SQL podoby pro oba kraje
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Kraj_CHR_HRA

Hradecky_kraj

SSMuller_hradec_vsechny_body_DB_cesky.stx - soubor obsahující sídla v Hradeckém kraji s českým popisem
SSMuller_hradec_vsechny_body_DB_cesky.sql - SQL podoba výše zmíněného souboru
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Naplneni_SIDLA_MULLER

SSMuller_hradec_vsechny_body_nem.stx - soubor obsahující sídla v Hradeckém kraji s německým popisem
SSMuller_hradec_vsechny_body_DB_cesky.sql - SQL podoba výše zmíněného souboru
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

HRA_typy_sidel

Soubory obsahující sídla dle jejich typu a jejich SQL podoby
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Chrudimsky_kraj

SSMuller_chrudim_vsechny_body_DB_cesky.stx - soubor obsahující sídla v Hradeckém kraji s českým popisem
SSMuller_chrudim_vsechny_body_DB_cesky.sql - SQL podoba výše zmíněného souboru
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Naplneni_SIDLA_MULLER

SSMuller_chrudim_vsechny_body_nem.stx - soubor obsahující sídla v Hradeckém kraji s německým popisem
SSMuller_chrudim_vsechny_body_DB_cesky.sql - SQL podoba výše zmíněného souboru
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

CHR_typy_sidel

Soubory obsahující sídla dle jejich typu a jejich SQL podoby
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Neexistující_sidla_CHR_HRA

hradec_dlouhy.txt - soubor s daty pro neexistující sídla v
Hradeckém kraji
hradec_dlouhy.sql - SQL podoba výše zmíněného souboru
chrudim_dlouhy.txt - soubor s daty pro neexistující sídla v
Chrudimském kraji
chrudim_dlouhy.sql - SQL podoba výše zmíněného souboru
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Kraj_LTM_RAK

SSMuller_LTM_JTSK.stx - data pro aktualizaci tabulky SIDLA_MULLER
SSMuller_RAK_JTSK.stx - data pro aktualizaci tabulky SIDLA_MULLER
LTM_sidla_muller.sql - správný formát dat pro aktualizaci
RAK_sidla_muller.sql - správný formát dat pro aktualizaci
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

LTM_typy_sidel

Soubory obsahující sídla dle jejich typu a jejich SQL podoby
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

RAK_typy_sidel

Soubory obsahující sídla dle jejich typu a jejich SQL podoby
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Práce_s_tabulkou_SIDLA

aktualizace_sidla_aktual

sidlaXaktual.stx - soubory pro aktualizaci tabulky SIDLA
sidla-updateX.sql - formát souborů pro vstup do DB
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

aktualizace_sidla_smaz

sidlaXsmaz.stx - soubory pro aktualizaci tabulky SIDLA
sidla-updateXsmaz.sql - formát souborů pro vstup do DB
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

sidla

sidla.stx - soubor s prvotními vstupními daty do tabulky SIDLA
poskytnutá ČSÚ
sidla_insert.sql - forma výše uvedeného souboru pro vložení dat do DB
Rozdeleni.java, Sidlo.java - program pro práci s daty v adresáři

Hájek_BP_2009.pdf

samotná bakalářská práce ve formátu .pdf

Příloha C: Přehled vzniklých tabulek v databázi sídel

| Název tabulky | Její určení |
|--------------------------------|---|
| SIDLA | Tabulka reprezentující současný stav sídel v ČR, v rámci jejich názvu, polohy a určení nynějšího typu sídla |
| SIDLA_MULLER | Tabulka reprezentující stav sídel na Müllerových rukopisných mapách na počátku 18. století, v rámci jejich německého názvu, polohy a určení tehdejšího typu sídla |
| NALEZITOST | Významová tabulka pro popis hodnoty sloupce zmeneno tabulky SIDLA |
| KRAJ | Referenční tabulka obsahující popis kraje, kde se sídlo vyskytovalo v rámci Müllerových rukopisných map a slouží k tomu, aby se do sloupce kraj tabulky SIDLA_MULLER neuložila nepovolená hodnota |
| TYP_SIDEL | Referenční tabulka obsahující popis typu sídla na Müllerových rukopisných mapách a slouží k tomu, aby se do sloupce typ tabulky SIDLA_MULLER neuložila nepovolená hodnota |
| ROZDIL_SIDLA_MULLER | Tabulka uchovávající údaje o rozdílu souřadnic a názvů odpovídajících si sídel mezi tabulkami SIDLA a SIDLA_MULLER |
| ROZDIL_SIDLA_MULLER_UNIKATNI | Tabulka uchovávající údaje o rozdílu souřadnic a odchylek odpovídajících si sídel s kardinalitou 1:1 mezi tabulkami SIDLA a SIDLA_MULLER |
| ROZDIL_SIDLA_MULLER_DUPLICITNI | Tabulka uchovávající údaje o rozdílu souřadnic a odchylek odpovídajících si sídel s kardinalitou 1:N mezi tabulkami SIDLA a SIDLA_MULLER |
| ODCHYLKY_SIDLA_MULLER | Tabulka s údaji o odchylkách odpovídajících si sídel (nebo průměru odchylek pro duplicitní data) mezi tabulkami SIDLA a SIDLA_MULLER |
| POCET_SHODNYCH_SIDEL | Tabulka znázorňující duplicitní sídla a počet těchto duplicit v tabulce SIDLA_MULLER |
| ROZDIL_V_MULLEROVI | Obsahuje rozdíly mezi souřadnicemi a německými názvy duplicitních sídel v tabulce SIDLA_MULLER |
| ROZDIL_V_MULLEROVI_UNIKATNE | Obsahuje rozdíly mezi souřadnicemi a polohové odchylky duplicitních sídel v tabulce SIDLA_MULLER, avšak v počtu jednoho záznamu pro každé takové sídlo |
| SHODNE_NAZVY | Uchovává ty záznamy sídel, které mají stejný český název v tabulce SIDLA jako německý v tabulce SIDLA_MULLER |
| VYBER_JINYCH_NEMECKYCH | Zobrazuje ta duplicitní sídla, která mají v každém svém záznamu v tabulce SIDLA_MULLER jiný německý název |
| MULLER_UNIQUE_ZAZNAMY | Obsahuje unikátně uložená sídla v tabulce SIDLA_MULLER, která v dnešní době stále existují |
| POMER_UNIQUE | Zobrazuje zastoupení sídel v tabulce MULLER_UNIQUE_ZAZNAMY v rámci typu těchto sídel vyobrazeným na Müllerových mapách |
| MULLER_REDUNDANTNI_ZAZNAMY | Obsahuje redundantně uložená sídla v tabulce SIDLA_MULLER, které v dnešní době stále existují |
| POMER_REDUNDANTNI | Zobrazuje zastoupení sídel v tabulce MULLER_REDUNDANTNI_ZAZNAMY v rámci jejich typu vyobrazeným na Müllerových mapách |
| MULLER_UNIQUE_I_REDUNDANTNI | Spojení tabulek MULLER_REDUNDANTNI_ZAZNAMY a MULLER_UNIQUE_ZAZNAMY |
| POMER_VSECH | Zobrazuje zastoupení sídel v tabulce MULLER_UNIQUE_I_REDUNDANTNI v rámci typu těchto sídel vyobrazeným na Müllerových mapách |
| MULLER_ZANIKLA_SIDLA | Obsahuje nyní již neexistující sídla s určením jejich typu dle klasifikace v 18. století |
| POMER_ZANIKLA | Zobrazuje zastoupení sídel v tabulce MULLER_ZANIKLA_SIDLA v rámci typu těchto sídel vyobrazeným na Müllerových mapách |
| VYSOKA_ID | Tabulka obsahující pouze dnes již neexistující sídla |

| Název tabulky | Její určení |
|----------------|--|
| TYP_CSU | Referenční tabulka obsahující popis typu sídla používaný v dnešní době a slouží k tomu, aby se do sloupce csu_typ tabulky SIDLA neuložila nepovolená hodnota |
| SROVNANI_OBCE | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení obce |
| SROVNANI_KU | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení definiční bod katastrálního území |
| SROVNANI_MCAST | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení městská část |
| SROVNANI_COBE | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení část obce evidenční |
| SROVNANI_COB | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení část obce |
| SROVNANI_ZSJ | Přehled počtu a typu sídel tabulky SIDLA_MULLER, kterým v dnešní době odpovídá označení základní sídelní jednotka |

Použitá literatura

- [1] *Demografický informační portál O DEMOGRAFII* [online], citováno [2009-03-28]
dostupný z URL <<http://demografie.info>>
- [2] Králíčková, P.: *Lokalizace rukopisných Müllerových map Prácheňského a Berounského kraje - bakalářská práce* [online],
vedoucí práce: Doc. Ing. Václav Čada, CSc., Plzeň, 2008
dostupná z Online archivu závěrečných prací studentů, URL
<<http://gis.zcu.cz/studium/dp/2008/>>
- [3] Müller, S.: *Lokalizace rukopisných Müllerových map Litoměřického a Rakovnického kraje - bakalářská práce* [online], vedoucí práce: Doc. Ing. Václav Čada, CSc.,
Plzeň, 2008
dostupná z Online archivu závěrečných prací studentů, URL
<<http://gis.zcu.cz/studium/dp/2008/>>
- [4] *Český statistický úřad | ČSÚ* [online], citováno [2009-01-21]
dostupný z URL <<http://www.czso.cz>>
- [5] *Licence PostgreSQL* [online], citováno [2009-01-17]
dostupné z URL
<http://www.pgsql.cz/index.php/1.03_Pod_jakou_licenc%C3%AD_je_PostgreSQL>
- [6] *charakteristika pgAdmin* [online], citováno [2009-01-17]
dostupné z URL <<http://www.pgadmin.org/>>
- [7] *SQL String Functions and Operators* [online], citováno [2009-05-09]
dostupné z URL <<http://www.postgresql.org/docs/8.3/static/functions-string.html>>
- [8] *Český statistický úřad | ČSÚ* [online] - *definice ZSJ*, citováno [2009-05-15]
dostupné z URL <http://www.czso.cz/sldb/sldb.nsf/i/zakladni_sidelni_jednotka>
- [9] Čada V., Vichrová M.: *Rukopisné mapy Čech J. Ch. Müllera*, Kartografické listy,
2009, 17.