

Úvod do zpracování prostorových dat

Kvalita bydlení ve vybraných částech Prahy

Zimní semestr 2018/2019

Tereza Kulovaná
Markéta Pecenová

Obsah

| | | |
|----------|-----------------------------------|-----------|
| 1 | Zadání | 2 |
| 1.1 | Zadání projektu | 2 |
| 1.2 | Zvolené téma | 2 |
| 2 | Data | 3 |
| 2.1 | Data a tématické vrstvy | 3 |
| 2.1.1 | ruain_praha | 3 |
| 2.1.2 | Otevřená data | 3 |
| 2.1.3 | Geoportál Praha | 3 |
| 3 | Zpracování dat | 4 |
| 3.1 | ruian_praha | 4 |
| 3.2 | Otevřená data | 4 |
| 3.3 | Geoportál Praha | 7 |
| 4 | Ukázka SQL dotazů | 9 |
| 5 | Závěr | 10 |
| 6 | Přílohy | 11 |
| 7 | Zdroje | 12 |

1 Zadání

1.1 Zadání projektu

Navrhňte a vytvořte tématické vrstvy (např. vodní toky, vodní plochy, lesy, silnice, železnice apod.) na základě dat *OpenStreetMap* a další otevřených zdrojů. Aplikujte testy datové integrity a odstraňte případné nekonzistence v datech. Vytvořte tutoriál - tj. sadu atributových a prostorových dotazů nad databází **pgis_uzpd**.

1.2 Zvolené téma

Jako téma pro semestrální projekt byla zvolena analýza kvality bydlení v Praze. Jelikož je toto téma velmi obecné a obsáhlé, bylo nutné zvolit užší zaměření. Z mnoha podkladů, které byly k dispozici, byly vybrány datové vrstvy, které nějak souvisely s občanskou vybaveností nebo zdravím obyvatel Prahy. Výsledným produktem je sada atributových a prostorových dotazů, které vyhodnocují kvalitu místa pro život na základě zvolených vstupních ukazatelů kvality.

2 Data

V rámci projektu bylo nutné získat potřebná vstupní data, nahrát je do databáze a zajistit jejich konzistenci. Nad těmito daty pak byly následně provedeny prostorové dotazy.

2.1 Data a tématické vrstvy

Data použitá v rámci projektu pochází ze tří zdrojů: ze schématu *ruian_praha*, portálů *Otevřená data* a *Geoportál Praha*.

2.1.1 ruian_praha

Schéma *ruian_praha* je součástí databáze **pgis_uzpd** a je v souřadnicovém systému JTSK.

Vrstvy:

- *adresni_mista* (adresnimista)
- *obvody* (spravniobvody)

2.1.2 Otevřená data

Veškeré datové vrstvy, které byly staženy z portálu *Otevřená data*, mají uvedené jako poskytovatele *HLAVNÍ MĚSTO PRAHA* a jsou vztažena pouze na území Prahy. Geometrie všech vrstev je reprezentována bodem. Data byla stažena ve formátu geoJSON a až na poslední uvedenou vrstvu byla v souřadnicovém systému WGS84 (S-JTSK nebyl k dispozici). Pouze *Vstupy do metra* byly staženy přímo v S-JTSK.

Vrstvy:

- *detska_hriste* (Dětská hřiště Praha, [Zdroj])
- *zdrav_zarizeni* (Lékárny a zdravotnická zařízení v Praze, [Zdroj])
- *odpad* (Mapa košů na tříděný odpad, [Zdroj])
- *metro* (Vstupy do metra, [Zdroj])

2.1.3 Geoportál Praha

Ze stránek *Geoportál Praha* byla stažena tématická vrstva zachycující zaplavené území Prahy při povodních v roce 2013. Geometrie této vrstvy je polygon a data byla stažena ve formátu geoJSON v systému JTSK. K dispozici ke stažení bylo vrstev mnohem více, například záplavové čáry pro stoletou/padesátiletou/dvacetiletou vodu atd. Pro zjednodušení byla stažena data, který byla ucelená a nejaktuálnější.

Vrstvy:

- *zaplavy2013* (Záplavové území 2013, [Zdroj])

3 Zpracování dat

3.1 ruian_praha

Adresní místa

Data byla stažena ze schématu *ruian_praha* (vrstva *adresnimista*) do nově vytvořené tabulky *adresni_mista*. Z původních dat byly zkopírovány pouze sloupce *kód*, *číslo popisné*, *číslo orientační* a sloupec s geometrií *geom*.

```
1 CREATE TABLE adresni_mista AS
2     SELECT kod, cislodomovni, cisloorientacni, geom
3     FROM ruian_praha.adresnimista
```

Následně byly nad sloupkem *kód* nastaven primární klíč a nad sloupkem *geom* prostorový index.

```
1 ALTER TABLE adresni_mista ADD PRIMARY KEY(kod)
2
3 CREATE INDEX adresy_index ON adresni_mista(geom)
```

Na závěr byla provedena kontrola validity geometrie, která vyšla negativní.

```
1 SELECT kod FROM adresni_mista WHERE NOT st_isvalid(geom)
```

Při zběžném prohlédnutí již zpracovaných dat bylo zjištěno, že výše uvedená funkce na kontrolu validity nefunguje zcela správně, jelikož neodstranila záznamy které měly ve sloupečku pro geometrii uvedenou hodnotu *NULL*. Bylo tedy nutné provést dodatečné pročištění dat níže uvedeným příkazem:

```
1 DELETE FROM adresni_mista WHERE geom IS NULL
```

Správní obvody Prahy

Data byla stažena ze schématu *ruian_praha* (vrstva *spravniobvody*) do nově vytvořené tabulky *obvody*. Geometrie dat je *multipolygon* a představují jednotlivé pražské správní obvody. Z původních dat byly zkopírovány pouze sloupce *ogc_fid*, *název* a sloupec s geometrií *geom*.

```
1 CREATE TABLE obvody AS
2     SELECT ogc_fid, nazev, geom
3     FROM ruian_praha.spravniobvody
```

Následně byly nad sloupkem *ogc_fid* nastaven primární klíč.

```
1 ALTER TABLE obvody ADD PRIMARY KEY(ogc_fid)
```

Na závěr byla provedena kontrola validity geometrie, která vyšla negativní.

```
1 SELECT nazev FROM obvody WHERE NOT st_isvalid(geom)
```

3.2 Otevřená data

Všechna data z tohoto zdroje byla stažena ve formátu geoJSON a převážně v systému WGS84 (výjimka: vrstva *metro* bylo od počátku v S-JTSK). Postup při zpracování těchto dat byl obdobný. Data byla nahrána do příslušného schématu pomocí nástroje *ogr2ogr*, a přímo v databázi byly daným vrstvám promazány nepotřebné sloupce a přejmenován

sloupeček s geometrií, který po nahrání dávkou získal uživatelsky nepřívětivý název *wkb_geometry*.

Jelikož se geometrie všech vrstev byla *bod*, pro kontrolu validity byla pro všechny použita stejná funkce:

```
1 SELECT ogc_fid FROM nazev_tabulky WHERE NOT st_isvalid (geom)
```

Dětská hřiště Praha

Během importu do databáze bylo nutné provést transformaci dat ze systému WGS84 (EPSG: 4326) do S-JTSK (EPSG: 5514). Při importu byl automaticky vytvořen sloupeček s primárním klíčem *ogc_fid* a byly vytvořeny prostorové indexy. Níže je uveden kód pro nahrání souboru do příslušného schématu databáze:

```
1 ogr2ogr -f "PostgreSQL" PG:"dbname=pgis_uzpd_user=uzpd18_d_password=
  d_uzpd18_host=geo102.fsv.cvut.cz" -t_srs 'EPSG:5514' "/mnt/c/Users/Terka
  /Documents/CVUT5-1/UZPD/projekt/detska_hriste.json" -nln uzpd18.d.
  detska_hriste
```

Příkaz pro smazání sloupečků, které pro další práci s daty nebyly nutné:

```
1 ALTER TABLE detska_hriste
2     DROP COLUMN url ,
3     DROP COLUMN name ,
4     DROP COLUMN perex ,
5     DROP COLUMN content ,
6     DROP COLUMN address ,
7     DROP COLUMN properties ,
8     DROP COLUMN image
```

Přejmenování sloupečku s geometrií:

```
1 ALTER TABLE detska_hriste
2     RENAME wkb_geometry TO geom
```

Na závěr byla provedena kontrola validity geometrie. Výsledná tabulka má tyto sloupce:

| detska_hriste | | | |
|---------------|------|----|----------|
| ogc_fid | geom | id | district |

Lékárny a zdravotnická zařízení v Praze

Během importu do databáze bylo nutné provést transformaci dat ze systému WGS84 (EPSG: 4326) do S-JTSK (EPSG: 5514). Při importu byl automaticky vytvořen sloupeček s primárním klíčem *ogc_fid* a byly vytvořeny prostorové indexy. Níže je uveden kód pro nahrání souboru do příslušného schématu databáze:

```
1 ogr2ogr -f "PostgreSQL" PG:"dbname=pgis_uzpd_user=uzpd18_d_password=
  d_uzpd18_host=geo102.fsv.cvut.cz" -t_srs 'EPSG:5514' "/mnt/c/Users/Terka
  /Documents/CVUT5-1/UZPD/projekt/zdrav_zarizeni.json" -nln uzpd18.d.
  zdrav_zarizeni
```

Příkaz pro smazání sloupečků, které pro další práci s daty nebyly nutné:

```

1 ALTER TABLE zdrav_zarizeni
2     DROP COLUMN id ,
3     DROP COLUMN address ,
4     DROP COLUMN email ,
5     DROP COLUMN web ,
6     DROP COLUMN telephone ,
7     DROP COLUMN opening_hours

```

Přejmenování sloupečku s geometrií:

```

1 ALTER TABLE zdrav_zarizeni
2     RENAME wkb_geometry TO geom

```

Na závěr byla provedena kontrola validity geometrie. Výsledná tabulka má tyto sloupce:

| zdrav_zarizeni | | | | |
|----------------|------|------|------|----------|
| ogc_fid | geom | name | type | district |

Mapa košů na tříděný odpad

Během importu do databáze bylo nutné provést transformaci dat ze systému WGS84 (EPSG: 4326) do S-JTSK (EPSG: 5514). Při importu byl automaticky vytvořen sloupeček s primárním klíčem *ogc_fid* a byly vytvořeny prostorové indexy. Níže je uveden kód pro nahrání souboru do příslušného schématu databáze:

```

1 ogr2ogr -f "PostgreSQL" PG:"dbname=pgis_uzpd_user=uzpd18_d_password=
  d_uzpd18_host=geo102.fsv.cvut.cz" -t_srs 'EPSG:5514' "/mnt/c/Users/Terka
  /Documents/CVUT5-1/UZPD/projekt/odpad.json" -nln uzpd18_d.odpad

```

Příkaz pro smazání sloupečků, které pro další práci s daty nebyly nutné:

```

1 ALTER TABLE odpad
2     DROP COLUMN objectid ,
3     DROP COLUMN id ,
4     DROP COLUMN stationnumber ,
5     DROP COLUMN stationname ,
6     DROP COLUMN citydistrictruiancode

```

Přejmenování sloupečku s geometrií:

```

1 ALTER TABLE odpad
2     RENAME wkb_geometry TO geom

```

Na závěr byla provedena kontrola validity geometrie. Výsledná tabulka má tyto sloupce:

| odpad | | | |
|---------|------|--------------|---------|
| ogc_fid | geom | citydistrict | pristup |

Vstupy do metra

Jelikož data již v S-JTSK byla, nebylo během importu nutné provést transformaci. Při importu byl automaticky vytvořen sloupeček s primárním klíčem *ogc_fid* a byly vytvořeny prostorové indexy. Níže je uveden kód pro nahrání souboru do příslušného schématu databáze:

```
1 ogr2ogr -f "PostgreSQL" PG:"dbname=pgis_uzpd_user=uzpd18_d_password=
  d_uzpd18_host=geo102.fsv.cvut.cz" "/mnt/c/Users/Terka/Documents/CVUT5-1/
  UZPD/projekt/metro.json" -nln uzpd18.d.metro
```

Příkaz pro smazání sloupečků, které pro další práci s daty nebyly nutné:

```
1 ALTER TABLE metro
2     DROP COLUMN objectid ,
3     DROP COLUMN vstupy_kod ,
4     DROP COLUMN vstupy_linka ,
5     DROP COLUMN vstupy_vest_kod ,
6     DROP COLUMN vstupy_vest_nazev ,
7     DROP COLUMN vstupy_vazba_bus ,
8     DROP COLUMN vstupy_vazba_csad ,
9     DROP COLUMN vstupy_vazba_kr ,
10    DROP COLUMN vstupy_vazba_pr ,
11    DROP COLUMN vstupy_vazba_privoz ,
12    DROP COLUMN vstupy_vazba_taxi ,
13    DROP COLUMN vstupy_vazba_tram ,
14    DROP COLUMN vstupy_vazba_vlak ,
15    DROP COLUMN zast_uzel_cislo ,
16    DROP COLUMN vstupy_popis ,
17    DROP COLUMN poskyt ,
18    DROP COLUMN vstupy_mimo_provoz
```

Přejmenování sloupečku s geometrií a názvu stanic:

```
1 ALTER TABLE metro
2     RENAME wkb_geometry TO geom ,
3     RENAME vstupy_uzel_nazev TO nazev
```

Na závěr byla provedena kontrola validity geometrie. Výsledná tabulka má tyto sloupce:

| metro | | |
|---------|------|-------|
| ogc_fid | geom | nazev |

3.3 Geoportál Praha

Záplavové území 2013

Ze stránek portálu *Geoportál Praha* byla data stažena ve formátu GeoJSON v systému JTSK, takže nebylo nutné provádět jejich další transformaci. Data byla nahrána do příslušného schématu pomocí nástroje *ogr2ogr*. Při importu byl automaticky vytvořen sloupeček s primárním klíčem *ogc_fid* a byly vytvořeny prostorové indexy. Níže je uveden kód pro nahrání souboru do příslušného schématu databáze:

```
1 ogr2ogr -f "PostgreSQL" PG:"dbname=pgis_uzpd_user=uzpd18_d_password=
  d_uzpd18_host=geo102.fsv.cvut.cz" "/mnt/c/Users/Terka/Documents/CVUT5-1/
  UZPD/projekt/zaplava2013.json" -nln uzpd18.d.zaplava2013
```

Příkaz pro smazání sloupečků, které pro další práci s daty nebyly nutné:

```
1 ALTER TABLE zaplava2013
2     DROP COLUMN objectid ,
3     DROP COLUMN nazev ,
4     DROP COLUMN typ
```


Přejmenování sloupečku s geometrií:

```
1 ALTER TABLE zaplava2013
2   RENAME wkb_geometry TO geom
```

Opět byla potřeba provést validita geometrie (polygonů). Níže uvedeným příkazem byly zjištěny chyby v datech:

```
1 SELECT id, geom, st_isvalidreason(geom) FROM uzpd18_d.zaplava2013 WHERE NOT
   st_isvalid(geom)
```

Ukázalo se, že nevalidní byly 4 polygony z celkových 86. Jako příčina těchto chyb se ukázala tzv. *Ring-self Intersection* polygonů. Nevalidní polygony byly opraveny vytvořením *bufferu* o velikosti 0.

```
1 UPDATE uzpd18_d.zaplava2013 SET geom = st_buffer(geom, 0.0) WHERE NOT
   st_isvalid(geom)
```

Výsledná tabulka má tyto sloupce:

| zaplava2013 | | | |
|-------------|------|--------------|------------|
| ogc_fid | geom | shape_length | shape_area |

4 Ukázka SQL dotazů

5 Závěr

6 Přílohy

- Příloha č. 1: Prezentace (prezentace.pdf)
- Příloha č. 2: SQL dávka (davka.sql)

7 Zdroje

1. *Otevřená data* [online] [cit. 28. 1. 2019].
Dostupné z: <http://www.geoportalpraha.cz/>
2. *Datové sady – Národní katalog otevřených dat (NKOD)* [online] [cit. 28. 1. 2019].
Dostupné z: <https://data.gov.cz/>
3. *LaTeX/Source Code Listing* [online] [cit. 28. 1. 2019].
Dostupné z: <https://en.wikibooks.org/>