Slovenská technická univerzita v Bratislave

Fakulta informatiky a informačných technológií

PostGIS

Martin Schön, Bc.

AIS ID: 103121

E-mail: xschon@stuba.sk

GitHub repozitár: https://github.com/FIIT-DBS/zadanie-pdt-xSchon

Predmet: Pokročilé databázové technológie

Zimný semester 2022/2023

Martin Schön

PostGIS

1. Import dát	3
2. Slovenské kraje	3
3. Kraje podľa veľkosti	4
4. Polygón domu	5
5. Kraj môjho domu	7
6. Aktuálna poloha v planet_osm_point	7
7. Som doma?	8
8. Vzdialenosť od FIIT	10
9. QGIS kraje a dom	11
10. Súradnice centroidu najmenšieho okresu	13
11. Úseky ciest bližšie ako 10km od hranice Pezinok – Malacky	14
12. Číslo katastrálneho územia s najdlhším úsekom cesty v mojom okrese	18
13. Oblast' Okolie_Bratislavy	20
Záver	22

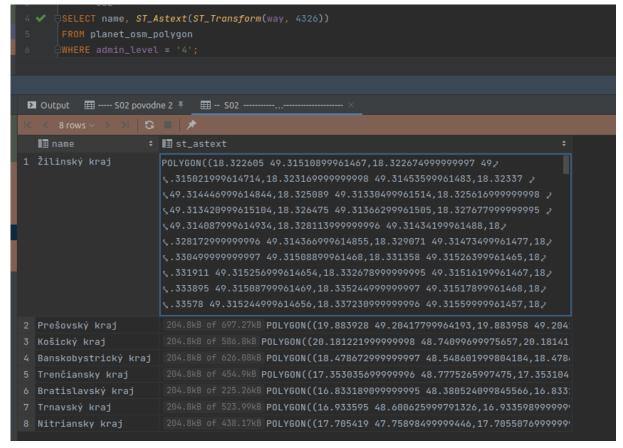
1. Import dát

Dáta som si stiahol z uvedenej stránky a importoval do novej, lokálnej databázy pomocou knižnice <u>osm2pqsql</u>. Na vypracovanie častí zadania, pri ktorých nebolo treba použiť QGIS som použil DataGrip s jeho predinštalovanou funkcionalitou "GeoViewer"-a, ktorá dokáže zobraziť súradnice pre vybrané súradnicové systémy.

2. Slovenské kraje

Niekde sa uvádza x,y ako x=longitude, y=latitude a niekde naopak - neexistuje jedno štandardizované poradie. Mení sa to medzi aplikáciami, zobrazeniami, je to rôzne na rôznych stránkach. Pri programovaní X zvyčajne predstavuje longitude a Y latitude (toto platí aj o nami použitých dátach). "Klasické" a teda dlhšie používané v geografii (dávno pred počítačmi) mapovanie ale bolo opačné. Preto napríklad aj na Google Maps X reprezentuje latitude a Y longitude. Toto predstavuje štandard pri zobrazovaní dát polohy na mapu - to znamená, že v počítači sú spravidla držané opačne, ako ich vo finále aplikácia zobrazuje. Pri prekreslovaní mojich výsledkov na mapu je teda (v priebehu celého tohto zadania) potrebné myslieť na túto skutočnosť.

V úlohe 2 máme transformovať všetky ways reprezentujúce hranice jednotlivých krajov do bodov. Vo finále mi tak ku každému kraju vznikne polygón bodov vo formáte WKT, teda vo formáte POLYGON ((x1 y1, x2 y2, ... xn yn, x1 y1)). Polygón musí byť uzavretý cyklus, preto sa začiatočný bod opakuje na konci. Každá dvojica x y predstavuje jeden bod v súradnicovom systéme 4326, kde (ako je vyššie spomenuté) x predstavuje longitude a y latitude bodu kraja.



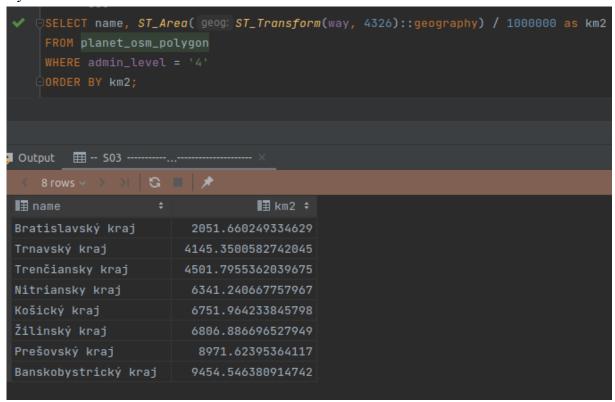
3. Kraje podľa veľkosti

Kraje vyberám rovnako, ako v úlohe 2, teda pomocou filtra hodnoty admin_level '4' z tabuľky *planet_osm_polygon*. Následne je tieto dáta potrebné pretransformovať z ich originálneho SRID 3857 na SRID 4326.



Keby po tejto transformácií hneď vykonám ST_Area a získam tak veľkosť, mohla by byť skreslená až o 40%, keďže sa jedná o mercatorovo zobrazenie (ktoré spôsobuje, že rovnaká plocha sa javí väčšia, čím je bližšie k pólu). Preto je potrebné transformovať tento výsledok na geography type, ktorý namapuje body na guľové zobrazenie, čím získame presnejšie dáta o veľkosti. Tie sú trochu náročnejšie na vyrátanie (keďže rátajú vzdialenosť na guli), no sú presné.

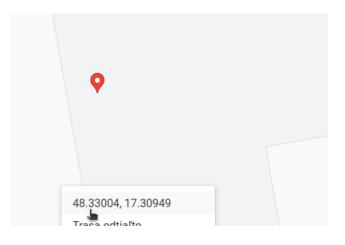
Finálny bod je predelenie veľkosti kraja miliónom, keďže prvotný výsledok je v m2 a my chceme dostať hodnotu v km2.



4. Polygón domu

Na získanie súradníc domu som našiel dve použiteľné metódy - prvou je export súradníc z oficiálnych dát <u>open street máp</u>. Je možné exportovať dáta a z nich získať súradnice z súradnicovom systéme 4326, teda ako longitude latitude. Alternatívou je získať tieto dáta z google máp - v rovnakom formáte, po jednotlivých bodoch.

Ja som sa rozhodol pre druhú možnosť, kedy som si vybral jednotlivé body na základe google máp a potom som tieto body preniesol do databázy.

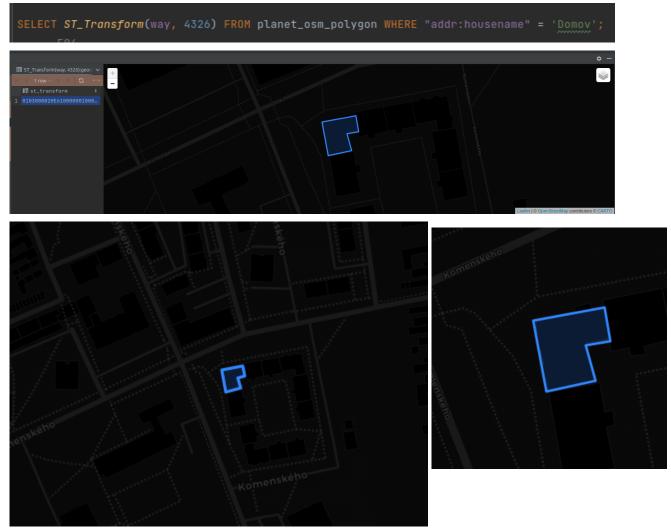


Prenesenie do databázy následne prebieha tak, že som z daných bodov vytvoril uzatvorený polygón. S pomocou funkcie ST_GeomFromText a WKT polygónu vznikol (v mojom prípade) polygón zložený zo 6tich bodov na súradniciach približne 17.30 a 48.33. Netreba zabúdať na to, že aj v tomto prípade PostGIS číta longitude latitude opačne, ako je to v prípade Google máp.

```
ST_Transform(ST_SetSRID( geog: ST_GeomFromText('POLYGON(( 17.309516058731177 48.3300009742030, 17.30967503789495 48.33002440991696, 17.30964503789495 48.330100328261695, 17.309725504163416 48.330110328261695, 17.309705504163416 48.330182479773576, 17.309471131731286 48.33015394805975, 17.309516058731177 48.3300009742030))'), srid: 4326), 3857));
```

Potom, čo som mal takto vytvorený polygón bolo potrebné uložiť ho v správnej súradnicovej sústave. V prípade way v databáze je to 3857. Samotný ST_Transform ale nefunguje priamo - pretože pri vytváraní polygónu cez ST_GeoFromText je súradnicová sústava nastavená na 0 - neznáma pre PostGIS. Preto som najskôr musel definovať, že sa jedná o polygón zo SRID 4326 a následne bola možná transformácia na formát vhodný pri vkladaní do DB - 3857. Celý import vyzerá takto (pridal som aj iné hodnoty pre ľahké hľadanie v DB):

Pri následnom vyhľadávaní súradníc domu tak môžem použiť tieto parametre. Pre vizualizáciu pomocou nástroja Geo Viewer v DataGripe je potrebné opäť transformovať way do súradnicového systému 4326. Potom je možné kliknutím na vrátený riadok vizualizovať polygón na mape



5. Kraj môjho domu

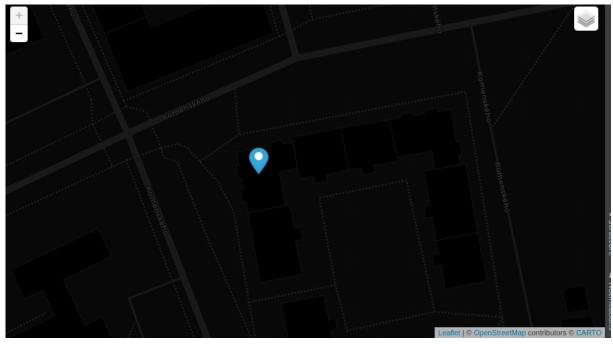
Na získanie názvu kraja, v ktorom sa nachádza môj dom som použil funkciu ST_Within. Tá vráti True, ak je jeden celý objekt súčasťou druhého objektu - čo v mojom prípade platí. Vybral som si teda všetky kraje, cross joinom vytvoril všetky dvojice dom–kraj a následne v nich našiel ten, v ktorom sa nachádzam.

6. Aktuálna poloha v planet_osm_point

Cez <u>SELECT ST_Srid(way) FROM planet_osm_point;</u> som si potvrdil, že aj v planet_osm_point je použitá súradnicová sústava 3857. To znamená, že postup vkladania do tabuľky databázy bude podobný, ako pri vkladaní polygónu mojej domácnosti - ibaže namiesto polygónu budem vkladať bod. Cez ST_GeometryType som zistil, že naozaj sú geometrické dáta v tabuľke planet_osm_point typu ST_Point.

Následne som pomocou Google máp našiel svoju polohu a vložil ju do databázy veľmi podobným štýlom, ako polygón v úlohe 4 - vložil som bod z WKT, nastavil mu súradnicovú sústavu 3857 a pretransformoval do 4326.

Zobrazenie na mape:

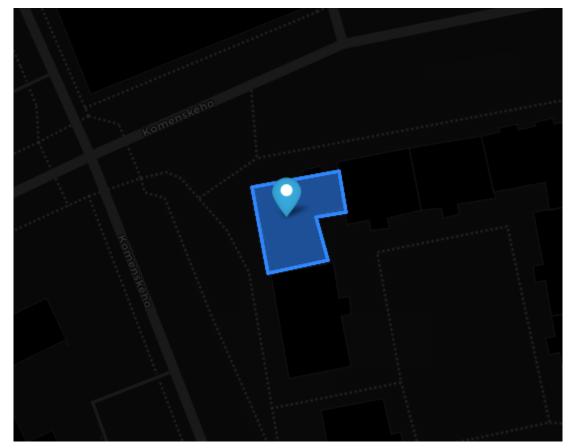


7. Som doma?

Pomocou CROSS JOINu som si vytiahol súradnice mojej aktuálnej polohy a namapoval ich na súradnice môjho domu. Cez ST_Within som následne zistil, či sa nachádzam doma - odpoveď na túto otázku je True, teda áno - nachádzam, keďže som lenivý a zadanie som začal vypracovávať až cez víkend :(.

Pre vizualizáciu v DataGripe je potrebné pri SELECTe transformovať home.way a me.way do súradnicového systému 4326.

```
SELECT ST_Transform(home.way, 4326) AS home,
ST_Transform(me.way, 4326) AS me,
```



8. Vzdialenosť od FIIT

```
73 

SELECT

74  ST_Distance(

75  geog1: ST_Transform(school.way, 4326)::geography,

76  geog2: ST_Transform(me.way, 4326)::geography) / 1000

AS distance_km

FROM planet_osm_polygon school

79  CROSS JOIN (SELECT way

80  FROM planet_osm_point

81  WHERE name = 'Martin Schon') AS me

82  WHERE name = 'Fakulta informatiky a informačných technológií STU';

SOUTPUT  SOS  CSV  ↑ ↑ ♠ €

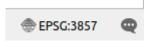
III distance_km 

1  26.31406597058
```

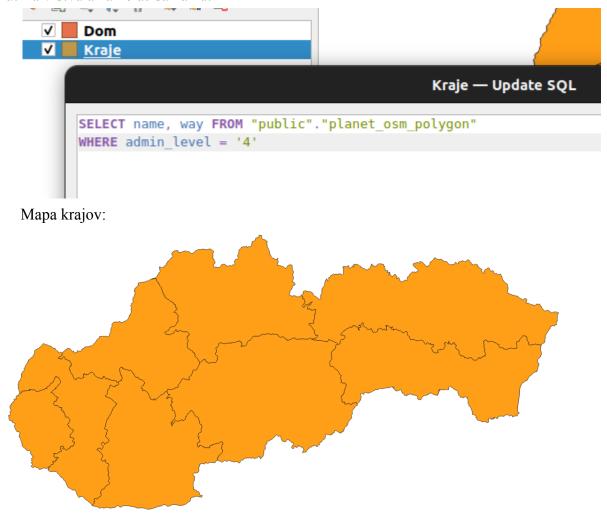
V tejto úlohe som využil vedomosti z predchádzajúcich úloh - CROSS JOINom som dostal svoj bod pri polygóne školy. Následne som pretransformoval dáta na geography, aby som získal reálnu vzdialenosť medzi nimi (podobne ako v úlohe 3). Vo finále som prerátal vzdialenosť na kilometre, aj keď to nebolo nevyhnutné - no ľahšie sa to číta.

9. QGIS kraje a dom

Stiahol som si QGIS, pripojil sa na lokálnu databázu cez DB Managera a dal som si vykresliť vrstvu všetkých polygónov. Potom som zistil, že si viem vytvárať a upravovať vrstvy, tak som pridal svoju vlastnú vrstvu reprezentujúcu kraje. Netreba vykonávať žiadnu transformáciu, je potrebné sa iba uistiť, že EPSG (v pravom dolnom rohu) je nastavené správne a reprezentuje SRID dát.



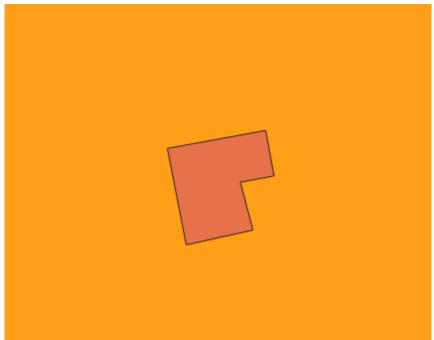
Na vykreslenie krajov som pridal vrstvu Kraje s jednoduchým SQL, ktoré ich vyberá ako v úlohe 2. Najväčší problém mi spôsobovalo to, že QGIS absolútne nezvláda, ak SQL príkaz končí bodkočiarkou, bez toho to ale mentálne ťažko prežívam ja. Inak nie je problém vykresliť si mapu krajov a obrázok. V prípade, ak sa stratím v prázdne je možné pravým kliknúť na vrstvu a zamerať sa na ňu.



Pri výbere môjho domu som použil query použitú aj v úlohe 4. Zvláštnosťou bolo, že QGIS si automaticky query upravil (ja som napísal iba tú v zátvorkách - SELECT way... Schon'). QGIS si moju query prečítal a zvládol ju vykresliť, ale pridal uid na unikátnu identifikáciu jednotlivých objektov (keďže môj dom neobsahuje id).



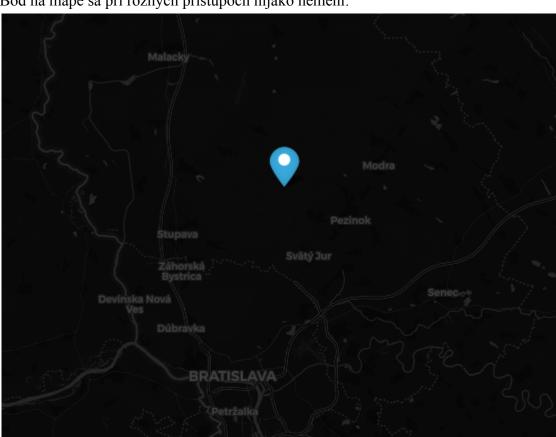
Následne som sa pokúsil nájsť svoj dom na mape Slovenska - je potrebné zoradiť layers v správnom poradí - konkrétne teda mať dom na vrchu krajov. To sa dá buď fyzickým potiahnutím vrstiev, alebo pravým klikom + Move To Top (prípadne move to bottom v prípade krajov). Každopádne potom som našiel dom na mape a tu je jeho screenshot:



10. Súradnice centroidu najmenšieho okresu

V tejto úlohe som opäť využil ST_Area, kedy som pomocou zoradenia ORDERom vybral najmenší kraj - Bratislavský. Následne som použil ST_Centroid funkciu, ktorá vyráta centroid polygónu. Tento výsledok som potom mohol transformovať na súradnicový systém 4326, z ktorého je možné vybrať ST_Y a ST_X reprezentujúci latitude a longitude súradnicu tohto bodu. Tým, že tieto dáta sú transformované, tak použitý SRID bol 4326.

Alternatívne by sa dalo použiť ST_AsText, ktoré vráti dáta vo formáte WKT Pointu, taktiež v 4326 SRID, nakoľko dáta boli transformované. Prípadne nemusím vôbec meniť dáta na AsText formát, ani transformovať - vždy dostanem centroid, hoci v poslednom prípade (bez transformácie) v SRID 3857 (takýto formát nemožno zobraziť na GeoViewere DataGripu, ale údaj je správny).



Bod na mape sa pri rôznych prístupoch nijako nemení:

11. Úseky ciest bližšie ako 10km od hranice Pezinok – Malacky

```
□CREATE TABLE roads_close_Pezinok_Malacky
□(id serial PRIMARY KEY,
osm_id BIGINT,
way geometry(LineString, 3857),
distance REAL,
name VARCHAR(255),
highway VARCHAR(255)
□);
```

Najskôr som si vytvoril tabuľku roads_close_Pezinok_Malacky, v ktorej si budem držať výsledok tejto operácie. Bolo samozrejme možné pridať aj viac stĺpcov, či vytvárať tabuľku priamo zo SELECTu, ale ja do nej iba vyberám dané hodnoty - nepotrebujem všetky stĺpce. Založil som si teda tabuľku s vybranými stĺpcami.

Následne túto tabuľku napĺňam cestami. Vnoreným SELECTom vyberiem priesek okresov Pezinok a Malacky - tým dostanem jednu líniu, ktorá reprezentuje hranicu (border) ktorá leží na hranici okresov Malacky a Pezinok. Používam ST_LineMerge, aby mi z hraníc vznikol pekný jeden objekt.

```
JSELECT border.border

JFROM (SELECT

ST_LineMerge(ST_Transform(ST_Intersection(malacky.way, pezinok.way), 4326)) as border

FROM planet_osm_polygon AS malacky,

(SELECT name, way

FROM planet_osm_polygon

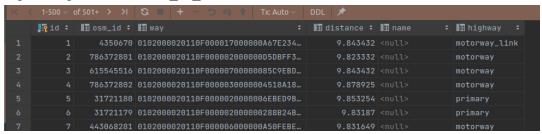
WHERE name LIKE 'okres Pezinok') AS pezinok

WHERE malacky.name LIKE 'okres Malacky') as border;
```

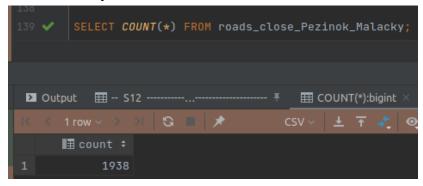


Následne z tabuľky planet_osm_roads vyberiem informácie o cestách. Keďže v tejto tabuľke sa nachádzajú aj iné čiary, ako tie, reprezentujúce cesty pre vozidlá, je potrebné ich filtrovať. To vykonávam pomocou stĺpca highway - ten v sebe drží informáciu o účele zapísanej geografickej hodnoty. V prípade osm existuje highway dokumentácia, ktorá opisuje význam jednotlivých zameraní. Stačilo mi teda vymazať niektoré čiary - pre chodcov, vodu,... Následne vymažem aj tie, ktoré majú highway rovné hodnote NULL, ale nemajú motorway rovné no (NULL je v pohode). Nakoniec vyberiem vzdialenosť medzi borderom a cestami pomocou ST_Distance nad 4326 geography cestami a vyberiem tie, ktoré sú bližšie, ako 10km.

Do databázovej tabuľky potom vložím údaje o vzdialenosti od hranice, linku cesty (v SRID 3857 na vloženie do databázy) a doplňujúce informácie. Súbor s výsledkami je možné nájsť v prílohe pod menom *z03 11 schon.csv* vo formáte:



Vo finále som dostal 1938 takýchto ciest.



Cesty vyzerajú nasledovne:



12. Číslo katastrálneho územia s najdlhším úsekom cesty v mojom okrese

Najskôr som si stiahol <u>dáta</u>, ktoré som si naimportoval do databázy pomocou command line príkazu ogr2ogr. Tie mi vytvorili v databáze 5 nových tabuliek: obec_0, okres_0, sr_0, kraj_0, ku_0. Pre mňa zaujímavá tabuľka je ku_0, keďže reprezentuje katastrálne územia. Katastrálne územia môjho okresu (Pezinok) vyzerajú takto:



V tejto úlohe som mal vybrať názov a id katastra, v ktorom sa nachádza najdlhší úsek cesty. Cesty v tomto prípade vyberám rovnako, ako v predchádzajúcom príklade - odstránim nemotoristické čiary z tabuľky planet_osm_roads. Okrem týchto ciest potrebujem získať informácie o katastroch - z novo vloženej tabuľky ku_0 vyberiem katastre nachádzajúce sa v Pezinskom okrese. O katastroch si držím informáciu o id a mene, z id5 a nm5 keďže sa hodnoty v nich zhodujú s oficiálnym zoznamom katastrov na Slovensku a teda dobre reprezentujú jednotlivé katastrálne územia.

Okrem toho si, samozrejme, vyberiem aj geografické znázornenie katastra, ktoré je potrebné pretransformovať, nakoľko je v SRID 5514. Potom môžem pomocou ST_Length vybrať dĺžku prieseku (ST_Intersection) jednotlivých ciest a katastrov. Z nich potom vyberiem najdlhší úsek. Používam aj podmienku ST_Intersects medzi cestami a katastrami, keďže to pomôže zrýchliť proces automatickým vyradením dvojíc ciest a katastrov, ktoré nie sú pri sebe. Vo finále mi tak na porovnanie ostanú len nenulové dĺžky, vyberiem cez ORDER BY a LIMIT najdlhší priesek. Potom vyberiem dĺžku a geografíu na vizualizáciu.

Výsledok, ktorý dostanem z mojej (vyššie ukázanej) query a jeho vizualizácia:





13. Oblasť Okolie Bratislavy

V tejto úlohe si najskôr vyberiem oblasť 20km okolo Bratislavy (od okraja mesta). Vykonám tak pomocou ST_Buffer -u, ktorý v SRID 4326 prijíma argument metrov (20000m), okolo ktorých má vytvoriť polygón oblasti okolo daného objektu (v mojom prípade okolo Bratislavy). Všetky polygóny v tejto úlohe použité z databázy pochádzajú z tabuľky planet osm polygon

```
Gelect ST_Transform(ST_Buffer(ST_Transform(way, 4326)::geography, 20000)::geometry, 3857) AS way
FROM planet_osm_polygon

WHERE name = 'Bratislava' AND admin_level = '6') AS okolie,
```



Následne odstránim z tohto polygónu samotnú Bratislavu - pomocou ST_Difference medzi okolím a polygónom Bratislavy. Tu je potrebné dbať na správne poradie polygónov.

```
SSELECT

ST_Transform(
ST_Difference( geom1: okolie.way, geom2: bratislava.way), 4326) AS Bratislava_Okolie

FROM

(SELECT ST_Transform(ST_Buffer(ST_Transform(way, 4326)::geography, 20000)::geometry, 3857) AS way

FROM planet_osm_polygon

WHERE name = 'Bratislava' AND admin_level = '6') AS okolie,

(SELECT way

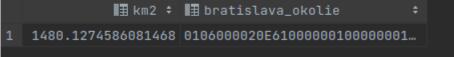
FROM planet_osm_polygon

WHERE name = 'Bratislava' AND admin_level = '6') AS bratislava;
```



A nakoniec vyberiem iba tú časť okolia, ktorá sa pretína so Slovenskom. Tú vyberiem z tabuľky a potom cez ST_Intersection vyberiem iba tú časť, ktorá sa nachádza na Slovensku v okolí Bratislavy. Takáto intersection mi vráti Collection, ktorú rozbalím a vyberiem z nej polygón (3). Z tohto výsledného polygónu potom získam jeho rozlohu v km2, ako aj jeho vizualizáciu.

Rozloha tejto oblasti je 1480 km2 a je znázornená na obrázku:





Záver

Zadanie som vypracoval samostatne, s využitím PostgreSQL verzie 15.0 a PostGIS verziou 3.3.1. V prílohe sa nachádza .csv súbor s výsledkami z úlohy 11, rovnako ako aj qgz súbor (spustiteľný v programe QGIS) z úlohy 10. Taktiež je priložený .sql súbor, v ktorom sa nachádzajú mnou použité SQL príkazy v ich finálnej verzii. Všetky tieto súbory je možné nájsť aj na <u>GitHube</u> uvedenom na titulnej strane. Ďakujem za čítanie dokumentácie mojej práce :).