

POSTGIS PSTGIS SPATIAL TRICKS

REGINA OBE



Ir@pcorp.us @reginaobe Consulting

Buy our books! at http://www.postgis.us/page_buy_book

BOOK IN PROGRESS: PGROUTING: A

PRACTICAL GUIDE

HTTP://LOCATEPRESS.COM/PGROUTING

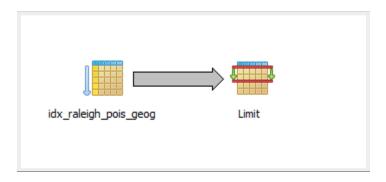


FIND N-CLOSEST PLACES (KNN)

Given a location, find the N-Closest places. Geography and n-D geometry operator support new in PostGIS 2.2.

EXAMPLE N-CLOSEST USING GEOGRAPHY DATA TYPE

Closest 5 Indian restaurants to here



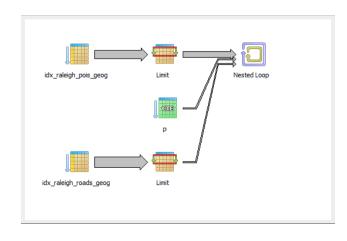
FIND 2 CLOSEST ROADS TO POINTS OF INTEREST WITH LATERAL

PostgreSQL 9.5+ and PostGIS 2.2+ for true distance.

```
WITH p AS (SELECT name, other tags->'amenity' As type,
ST Point(-78.64040,35.77627)::geography <-> geog As dist m, geog
FROM raleigh pois As pois
WHERE other tags @> 'cuisine=>indian'::hstore
ORDER BY dist m LIMIT 4)
SELECT p.name, p.type, r.name As road, r.dist m road::numeric(10,2), p.dist m::numeric(10,2) FROM p, LATERAL (SELECT rr.name, rr.geog <-> p.geog AS dist_m_road FROM raleigh roads AS rr WHERE rr.name > ''
ORDER BY dist m road LIMIT 2) As r;
                                                                                          | dist m road
                                    restaurant | West Lane Street
                                                                                                               1059.16
Blue Mango
                                                                                                  14.64
Blue Mango
                                  | restaurant | Glenwood Avenue
                                                                                                 16.62
                                                                                                             1059.16
                                                                                           111.25 | 19087.63
112.37 | 19087.63
Kadhambam Spices
                                                | Hatchet Creek Greenway
                                                | Hatchet Creek Greenway
Kadhambam Spices
```

: (8 rows) Time: 45.210 ms

Sitar



| restaurant | Chapel Hill Blvd Service Road | 36.62 | 35408.81

WHAT PLACES ARE WITHIN X-DISTANCE

Limit results set by distance rather than number of records. Like KNN, geometry can be anything like distance from a road, a lake, or a point of interest.

EXAMPLE: GEOGRAPHY WITHIN 1000 METERS OF LOCATION

What are closest fast food joints within 1000 meters. This will work for PostGIS 1.5+

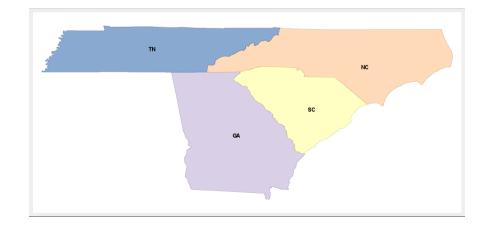
```
-- Time: 2.241 ms
SELECT name, other tags->'cuisine' As cuisine,
ST Distance(pois.geog, ref.geog) As dist_m
FROM raleigh pois AS pois,
(SELECT ST Point(-78.64040, 35.77627)::geography) As ref(geog)
   WHERE other tags @> 'amenity=>fast food'::hstore AND ST DWithin (pois.geog, ref.geog, 1000)
ORDER BY dist \overline{m};
Chick-fil-A
                                   | chicken | 115.15429719
                                 | sandwich | 208.9641767
Ouiznos
zpizza
                                 | pizza | 246.56944119
Snoopy's
                                                 | 851.75116195
Ouiznos Sandwich Restaurant | sandwich | 890.35270577
Char Grill
                         | burger | 906.69076176
Bruger's Bagels
                                                 1 997.0456652
(7 rows)
```

DIVIDE LARGE GEOMETRIES INTO SMALLER ONES WITH ST_SUBDIVIDE

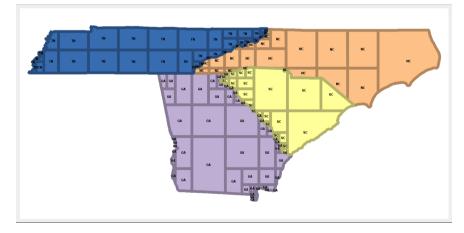
New in PostGIS 2.2. Works for non-point geometries (only 2D). Second arg is max number of points to allow per divide.

```
SELECT stusps, ST_SubDivide(geom, 1000) AS geom
FROM states
WHERE stusps IN('TN', 'NC', 'SC', 'GA');
```

Before had 4 rows



After have 186 rows

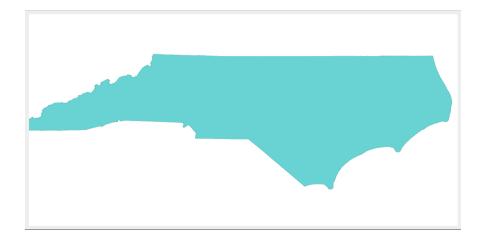


NEW POSTGIS 2.3 ST_GENERATEPOINTS

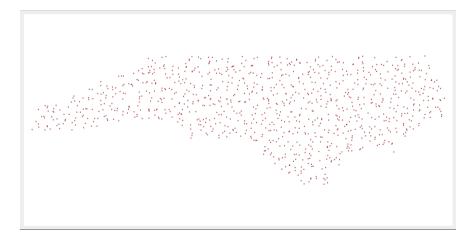
Converts multipolygon/polygon to multpoint - random space filling the area

```
SELECT stusps, ST_GeneratePoints(geom, 1000) AS geom
FROM states
WHERE stusps = 'NC';
```

Before: 1 Multi-Polygon



After: 1 multipoint of 1000 points



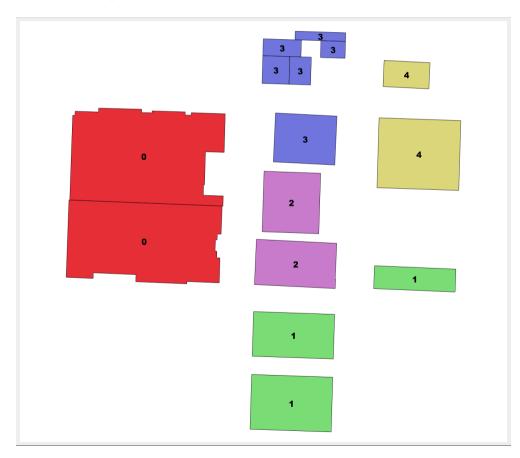
CLUSTERING GEOMETRIES USING WINDOW FUNCTIONS: COMING POSTGIS 2.3

- 2.3: ST_ClusterKMeans
- 2.3: ST_ClusterDbSCAN

ST_CLUSTERKMEANS: NUMBER BUILDINGS FROM 0-4 BY PROXIMITY TO EACH OTHER

```
SELECT name, ST ClusterKMeans(geom, 5) OVER() AS bucket
FROM raleigh poTys
WHERE name > '' AND building > ''
AND ST DWithin(geom, 'SRID=2264; POINT(2106664 737626)'::geometry, 500)
ORDER BY bucket;
```

Need to add geom column to view



name	bucket
Wake County Justice Center	l 0
Wake County Public Safety Center	0
Sir Walter Raleigh Hotel	1
Capital Bank Plaza	1
Sheraton Raleigh Hotel	1
Wake County Office Building	2
Wake County Courthouse	2
Kings] 3
The Mahler] 3
Capital Club 16] 3
CrossFit Invoke] 3
Federal Building] 3
North State Bank] 3
PNC Plaza	4
First Citizens Bank	4
(15 rows)	

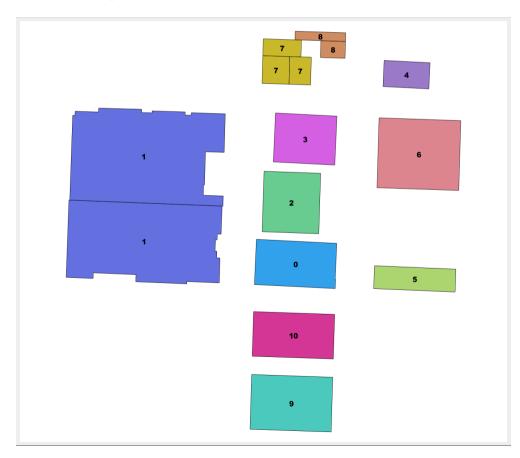
Time: 1.228 ms

ST_CLUSTERDBSCAN: SIMILAR TO KMEANS, BUT USES DESIRED DISTANCE AND MINIMUM NUMBER ITEMS

Cluster together buildings that intersect each other.

```
SELECT name, ST ClusterDBSCAN(geom, 0, 2) OVER() AS bucket, geom
FROM raleigh poTys
WHERE name > '' AND building > ''
AND ST DWithin(geom, 'SRID=2264; POINT(2106664 737626)'::geometry, 500)
ORDER BY bucket;
```

Need to add geom column to view



name	bucket
name	bucket 0 1 2 3 4 5 6 7 7 7
(15 rows)	10

Time: 1.046 ms

PARALLELIZATION OF SPATIAL JOINS AND FUNCTIONS

Will require PostgreSQL 9.6+ and PostGIS 2.3+. Read more: http://blog.cleverelephant.ca/2016/03/parallel-postgis-joins.html

Not yet committed to PostGIS repo, go here - https://github.com/pramsey/postgis/tree/parallel

```
set parallel tuple cost=0.01;
SET max parallel degree=2;
```

SEGMENTIZE A LINESTRING IN GEOGRAPHY

PostGIS 2.1+ ST_Segmentize(geography) can create great circles

GEOGRAPHY SEGMENTIZE VS. GEOMETRY SEGMENTIZE ON A MAP



From BoundlessGeo docs

SEGMENTIZE IN GEOGRAPHY OUTPUT AS GEOMETRY WKT

```
SELECT ST_AsText(
   ST_Segmentize('LINESTRING(-118.4079 33.9434, 2.5559 49.0083)'::geography,
   10000) );

LINESTRING(-118.4079 33.9434,-118.365191634689 33.9946750650617,
   -118.322351004015 34.0460320153076,
   ...,2.48756947085441 49.0516183725212,2.5559 49.0083)
```

SEGMENTIZE AND OUTPUT AS GOOGLE ENCODED LINE

PostGIS 2.2 we have ST_AsEncodedPolyline useful for drawing on google maps and use in Leaflet.

ST_LineFromEncodedPolyline for getting back a geometry.

ADDRESS STANDARDIZATION / GEOCODING / REVERSE GEOCODING

PostGIS 2.2 comes with extension address_standardizer. Also included since PostGIS 2.0 is postgis_tiger_geocoder (only useful for US).

In works improved address standardizer and worldly useful geocoder - refer to: https://github.com/woodbri/address-standardizer

ADDRESS STANDARDIZATION

Need to install address_standardizer, address_standardizer_data_us extensions (both packaged with PostGIS 2.2+). Using json to better show non-empty fields

```
SELECT *
FROM json each text(to json(standardize_address('us_lex', 'us_gaz', 'us_rules', '300 S. Salisbury St', 'Raleigh, NC 27601')))
WHERE value > '';

key | value
house_num | 300
predir | SOUTH
name | SALISBURY
suftype | STREET
city | RALEIGH
state | NORTH CAROLINA
postcode | 27601
(7 rows)
```

Same exercise using the packaged postgis_tiger_geocoder tables that standardize to abbreviated instead of full name

GEOCODING USING POSTGIS TIGER GEOCODER

Given a textual location, ascribe a longitude/latitude. Uses postgis_tiger_geocoder extension requires loading of US Census Tiger data.

```
SELECT pprint addy(addy) As address,
   ST X(geomout) AS lon, ST Y(geomout) As lat, rating
   FROM geocode('300 S. Salisbury St, Raleigh, NC 27601',1);

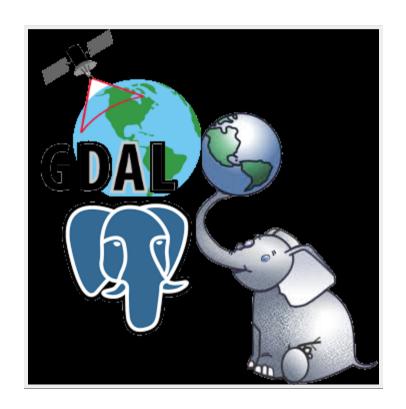
address | lon | lat | rating

300 S Salisbury St, Raleigh, NC 27601 | -78.6404024546499 | 35.7762672906178 | 0
(1 row)
```

REVERSE GEOCODING

Given a longitude/latitude or GeoHash, give a textual description of where that is. Using postgis_tiger_geocoder reverse_geocode function

GDAL CONJOINS WITH POSTGIS AND POSTGRESQL



- Scene 1: PostGIS Raster
- Scene 2: OGR_FDW Foreign Data Wrapper

SCENE 1: POSTGIS + GDAL = POSTGIS RASTER

A long time ago, a crazy man named Pierre Racine had a very crazy idea: https://trac.osgeo.org/postgis/wiki/WKTRaster and he got others Bborie Park, Sandro Santilli, Mateusz Loskot, David Zwarg and others to help implement his crazy scheme.

REGISTER YOUR RASTERS WITH THE DATABASE: OUT OF DB

You could with raster2pgsql the -R means just register, keep outside of database. Without the -R the data is stored in Db

```
raster2pgsql -I -C -R C:/data/nc_aerials/*.tif -F aerials | psql
```

OR (useful especially if you are on windows to force recursion of folders). Requires PostgreSQL 9.3+ PostGIS 2.1+

```
CREATE TABLE dir list(file name text);
COPY dir list FROM PROGRAM 'dir C:\data\nc_aerials\*.tif /b /S' WITH (format 'csv');
CREATE TABLE aerials ( rid serial PRIMARY KEY, rast raster, filename text);
INSERT INTO aerials(rast, filename)
SELECT
         ST AddBand(
                   NULL::raster,
                  d.file name, NULL::int[]
         ), d.file name
FROM dir list AS d;
SELECT AddRasterConstraints('aerials', 'rast');
--verify constraints
SELECT srid, scale x, scale y, blocksize x As width, blocksize y As height, pixeT_types, out_db FROM raster columns
WHERE r tabTe name = 'aerials';
srid | scale_x | scale_y | width | height | pixel_types | out_db
2264 | 0.5 | -0.5 | 10000 | 10000 | {8BUI,8BUI,8BUI} | {t,t,t}
(1 \text{ row})
CREATE INDEX idx aerials rast ON aerials USING gist(ST ConvexHull(rast));
analyze aerials;
```

LET'S TILE THE RASTER TO 200X200 CHUNKS RIGHT IN DB

Requires PostGIS 2.1+. ST_Tile, if working on out-db keeps out-db and very fast.

CREATE OVERVIEWS RIGHT IN DB

Requires PostGIS 2.2+. This will make in-db raster from outdb so might take a while. Took 8 minutes for my aerials table that had 30 10000x10000 raster links.

RETURN AN AREA: 500 FEET AROUND US

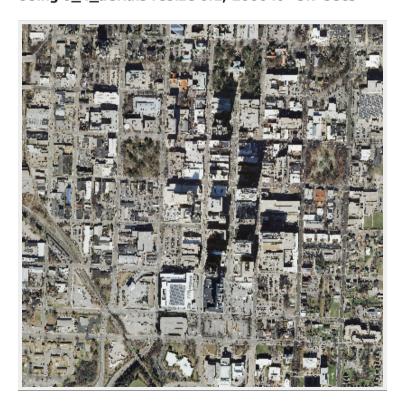
Project to same spatial ref as raster (2264 NC State Plane ft)

Using aerials: 4 secs (1 row), aerials_200_200: 5.9 sec (120 rows)

o_4_aerials resize 0.5 (980ms 1 row)



Using o 4 aerials resize 0.2, 2000 ft - 5.7 secs



SCENE 2: POSTGRESQL + GDAL ~ POSTGIS = OGR_FDW POSTGRESQL FOREIGN DATA WRAPPER

5 years ago I asked, https://trac.osgeo.org/postgis/ticket/974 and someone finally did it. It was slicker than I ever imagined.

DATA WRANGLING WITH OGR_FDW

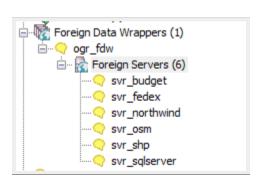
If you have all sorts of data of both a spatial and non-spatial flavor to tame, make sure you have ogr_fdw foreign data wrapper in your tool belt.

- For windows users, it's part of PostGIS 2.2 bundle on application stackbuilder.
- For CentOS/Red Hat/Scientific etc, it's available via yum.postgresql.org
- For others, if you have PostGIS with GDAL support, just need postgresql dev package to compile. Download the source https://github.com/pramsey/pgsql-ogr-fdw

WHY IS OGR_FDW SO SEXY?

You have the combined power of GDAL and PostgreSQL working seamlessly together. So many kinds of data you can query and take advantage of PostgreSQL functions and any extension functions and types such as PostGIS, hstore, built-in json.

- Spreadsheets
- ODBC datasources
- OSM files (OSM, PBF)
- ESRI Shapefiles
- Many more



ENABLE IT IN YOUR DATABASE

CREATE EXTENSION ogr_fdw;

LINK IN A WHOLE FOLDER OF ESRI SHAPEFILES AND DBASE FILES

QUERY YOUR SHAPE FILES LIKE REGULAR OLD TABLES

OGR_FDW NOW UPDATEABLE IF GDAL DRIVER ALLOWS WRITE

Version 1.0.1 brought IMPORT FOREIGN SCHEMA, latest in master branch supports updating, ability to include subset of columns, and detect srid

Check out the code and test: Download the source https://github.com/pramsey/pgsql-ogr-fdw

Windows users, winnie builds whenever master changes - http://postgis.net/windows_downloads/, look in extras folder for your PostgreSQL version - e.g 9.5 64-bit would be in pg9.5 extras folder and called ogrfdw-pg95-binaries-1.0w64gcc48.zip

DO AN UPDATE/INSERT/DELETE TO SHAPE FILE AS IF IT WERE A LOCAL TABLE

```
UPDATE shps.airports
SET geom = ST SnapToGrid(geom, 0.00001)
WHERE locid = 'JFK' RETURNING locid, ST AsText(geom) As wkt;
 locid | wkt
JFK | POINT(-73.77893 40.63975)
(1 \text{ row})
UPDATE 1
INSERT INTO shps.airports(locid, geom)
SELECT 'ROO', geom
FROM shps.airports
WHERE locid = 'BOS' RETURNING locid, ST_AsText(geom);
locid | st_astext
 ROO | POINT (-71.0064166666667 42.3629722222222)
INSERT 1
DELETE FROM shps.airports
WHERE locid = 'ROO';
```

OSM FILES

```
-- data from https://mapzen.com/data/metro-extracts/
CREATE SERVER svr osm
   FOREIGN DATA WRAPPER ogr fdw
  OPTIONS (datasource 'C:/fdw data/raleigh north-carolina.osm.pbf', format 'OSM');
  CREATE SCHEMA IF NOT EXISTS osm;
IMPORT FOREIGN SCHEMA ogr all
FROM SERVER svr osm INTÓ osm;
\dE osm.*
             List of relations
Schema | Name | Type | Owner
osm | lines | foreign table | postgres
osm | multilinestrings | foreign table | postgres
osm | multipolygons | foreign table | postgres osm | other relations | foreign table | postgres
    | points | foreign table | postgres
osm
(5 rows)
-- requires CREATE EXTENSION hstore;
CREATE TABLE raleigh pois AS
SELECT osm id, name, geom::geography As geog, is in,
    place, other tags::hstore
FROM osm.points;
CREATE TABLE raleigh roads AS
SELECT osm id, name, geom::geography As geog,
    other Tags::hstore
FROM osm. Tinés
WHERE highway > '';
```

LINK EVEN NON-SPATIAL LIKE MS ACCESS DATABASE TABLES AND QUERIES

```
CREATE SERVER svr northwind FOREIGN DATA WRAPPER ogr fdw
OPTIONS (datasource 'C:/fdw data/northwind.mdb',
 format 'ODBC'
CREATE SCHEMA IF NOT EXISTS northwind;
-- will link in all tables and queries starting with Ord
IMPORT FOREIGN SCHEMA "Ord"
FROM SERVER svr northwind INTO northwind;
\dE northind.*
                     List of relations
 Schema |
northwind | order details | foreign table | postgres
northwind | order details extended | foreign table | postgres
northwind | order subtotals | foreign table | postgres
northwind | orders
                                  | foreign table | postgres
northwind | orders gry
                                | foreign table | postgres
(5 rows)
```

The schema part is case sensitive, has to match source

EVEN SPREADSHEETS

Each workbook is considered a server and each sheet a table

Before

fid	Type		Zone A 2						
2 3	IntlPriority IntlPriority IntlPriority	0 -1	40.25	41.5 67.75	43 62.25	54.75	116.5 132	52 68	
4									>

After

type	weight	zone	price
IntlPriority IntlPriority IntlPriority IntlPriority	0 0	Zone A Zone B Zone C Zone D	40.25 41.5 43 54.75
•			

EVEN CSV FILES

You can point at a single CSV file or a whole folder of CSV files. Each file is considered a table.

Folder of CSV files

```
CREATE SERVER svr census FOREIGN DATA WRAPPER ogr_fdw
OPTIONS (datasource 'C:/fdw_data/census',
    format 'CSV'
);

IMPORT FOREIGN SCHEMA ogr all
FROM SERVER svr_census INTO public;
```

Single file

```
CREATE SERVER svr_census income FOREIGN DATA WRAPPER ogr_fdw
OPTIONS (datasource 'C:/fdw_data/census/income.csv',
   format 'CSV'
);

IMPORT FOREIGN SCHEMA ogr all
FROM SERVER svr_census_income INTO public;
```

EVEN OTHER RELATIONAL DATABASES

Format for SQL Server 'ODBC:your_user/your_password@yourDSN,table1,table2'. ODBC can be slow with a lot of tables (more than 150) so filter list if you have over 200 tables

FIN

BUY OUR BOOKS HTTP://WWW.POSTGIS.US