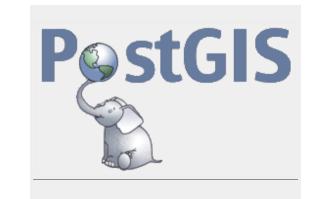


TEN PROBLEMS SOLVED BY POSTGIS



LEO HSU AND REGINA OBE

Presented at PGConfUS 2017

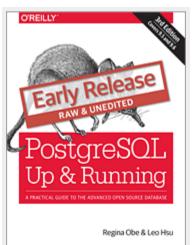
Buy our books! at http://www.postgis.us/page_buy_book

Books in progress: pgRouting: A Practical Guide PostgreSQL: Up and Running 3rd Edition









PROXIMITY ANALYSIS

N closest things. Things within x distance of this. Things that are within another. Both 2 and 3d geometries, 2d geodetic (aka geography), and even raster.

EXAMPLE N-CLOSEST USING GEOGRAPHY DATA TYPE

Closest 5 restaurants to here and kind of cuisine

GEOGRAPHY WITHIN 1000 METERS OF LOCATION

Works for geometry as well, but measurements and coordinates are in units of the geometry, not always meters.

name	type	cuisine	dist_m
Chilis Battello Starbucks Park and Sixth Torico Ice Cream The Kitchen at Grove Station Rustique	restaurant restaurant cafe restaurant fast_food restaurant restaurant	coffee_shop american	183.54545201 338.39714681 350.25322227 632.12204878 741.32599554 764.72996783 822.04122537
Helen's Pizza	restaurant		866.65681852

HOW MANY SUBWAY STOPS IN EACH BOROUGH?

```
SELECT b.boro name, COUNT(s.stop id) As num stops
FROM nyc boros AS b INNER JOIN nyc subways stops AS s ON ST_Covers(b.geom, s.geom)
GROUP BY b.boro name
ORDER BY b.boro_name;
```

boro_name		num_stops
	+-	
Bronx		70
Brooklyn		169
Manhattan		151
Queens		82
Staten Island		21

PROXIMITY WITH 3D DATA

If you have things like oil pipe lines and using linestrings with a Z component, it's just like ST_Distance, except you want to use ST_3DDistance, ST_3DDWithin, and ST_3DIntersects. These are part of the core postgis extension.

For more advanced 3d, like if you need ST_3DIntersection, and ST_3DIntersects that does true surface and solid analysis (PolyhedralSurfaces), you'll want to install extension postgis_sfcgal.

INTERSECT RASTER AND GEOMETRY: RASTER VALUE AT A GEOMETRIC POINT

```
SELECT pois.name, ST Value(e.rast,1,pois.geom) AS elev
FROM pois INNER JOIN nj ned As e ON ST_Intersects(pois.geom,e.rast)
WHERE pois.tags ? 'cuisine'
ORDER BY ST_SetSRID(ST_Point(-74.036,40.724),4269) <-> pois.geom
LIMIT 5;
```

name	6	elev
	-+	
Chilis	2	2.64900875091553
Starbucks	2	2.61004424095154
Battello	2	2.18213820457458
Park and Sixth	3	3.79218482971191
The Kitchen at Grove Station	2	2.06850671768188

REPROJECT ON-THE-FLY

DATABASE COLUMN TYPE TRANSFORMATION AND CONVERSION FOR GEOMETRY AND GEOGRAPHY

Convert from current projection to NYC State Plane feet (look in spatial_ref_sys for options).

```
ALTER TABLE nyc boros
ALTER COLUMN geom TYPE geometry(Multipolygon, 2263)
USING ST_Transform(geom, 2263);
```

Convert geometry to geography

```
ALTER TABLE nyc boros
ALTER COLUMN geom TYPE geography (Multipolygon, 4326)
USING ST_Transform (geom, 4326)::geography;
```

Convert back to geometry

```
ALTER TABLE nyc boros
ALTER COLUMN geom TYPE geometry(Multipolygon, 2263)
USING ST Transform(geom::geometry,2263);
```

ST_TRANSFORM FOR RASTER

For more info, read the manual http://postgis.net/docs/RT_ST_Transform.html. The algorithm defaults to NearestNeighbor algorithm, fastest but not the most appealing

```
SELECT ST_Transform(rast,3424) AS rast
FROM nj_ned
WHERE ST_Intersects(rast,ST_SetSRID(ST_Point(-74.036,40.724),4269));
```

You can override the warping algorithm

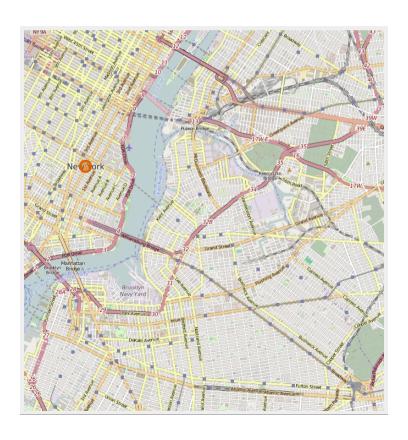
```
SELECT ST_Transform(rast,3424,'Lanczos') AS rast
FROM nj_ned
WHERE ST_Intersects(rast,ST_SetSRID(ST_Point(-74.036,40.724),4269));
```

Creating a whole new transformed table, align your rasters. This ensures rasters have same grid and pixel size.

```
WITH a AS (SELECT ST Transform(rast, 3424, 'Lanczos') AS rast FROM nj ned LIMIT 1) SELECT Fid, ST Transform(n.rast,a.rast,'Lanczos') AS rast INTO nj ned 3424 FROM nj ned AS n, a;
```

3. MAP TILE GENERATION

Common favorite for generation tiles from OpenStreetMap data. Check out TileMill and MapNik which both read PostGIS vector data and can generate tiles. Various loaders to get that OSM data into your PostGIS database: osm2pgsql, imposm, GDAL. TileMill is a desktop tool and MapNik is a toolkit with with python bindings and other language bindings.



OUTPUT SPATIAL DATA IN MANY FORMATS

GeoJSON, KML, SVG, and TWB (a new light-weight binary form in PostGIS 2.2). Coming in PostGIS 2.4 is ST_AsMVT (for loading data in MapBox Vector Tiles format) GeoJSON commonly used with Javascript Map frameworks like OpenLayers and Leaflet.

3D VISUALIZATION

X3D useful for rendering PolyhedralSurfaces and Triangular Irregulated Networks (TINS), PolyHedralSufaces for things like buildings. TINS for Terrain

Checkout https://github.com/robe2/node_postgis_express built using NodeJS and http://www.x3dom.org (X3D in html 5)

3D PROXIMITY AND RENDERING

Use 3D bounding box &&& operator and form a 3D box filter

X3Dom with texture



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 also to show fields

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, FROM geocode(*29 Fort Greene Pl, Brookl\overline{y}n, NY 11217',1);
```

address	lon	lat	rating
29 Fort Greene Pl, New York, NY 11217	-73.976819945824	40.6889624828967	8

REVERSE GEOCODING

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

PHOTOSHOP WITH POSTGIS

Pictures are rasters. Rasters are pictures. You can manipulate them en masse using the power of PostGIS raster.

READING PICTURES STORED OUTSIDE OF THE DATABASE: REQUIREMENT

new in 2.2 GUCS generally set on DATABASE or system level using ALTER DATABASE SET or ALTER SYSTEM. In PostGIS 2.1 and 2.0 needed to set these as Server environment variables.

```
SET postgis.enable outdb rasters TO true;
SET postgis.gdal_enabled_drivers TO 'GTiff PNG JPEG';
```

REGISTER YOUR PICTURES WITH THE DATABASE: OUT OF DB

You could with raster2pgsql the -R means just register, keep outside of database:

```
raster2pgsql -R /data/Dogs/*.jpg -F pics | psql
```

OR

```
CREATE TABLE pics (file_path text);
COPY pics FROM PROGRAM 'ls /data/Dogs/*.jpg';
ALTER TABLE pics ADD COLUMN rast raster;
ALTER TABLE pics ADD COLUMN file_name text;

-- Update record to store reference to picture as raster, and file_name
UPDATE pics SET rast = ST_AddBand(NULL::raster, file_path, NULL::int[]),
    file_name = split_part(file_path,'/',4);
```

GET BASIC RASTER STATS

This will give width and height in pixels and the number of bands. These have 3 bands corresponding to RGB channels of image.

RESIZE THEM AND DUMP THEM BACK OUT

This uses PostgreSQL large object support for exporting. Each picture will result in a picture 25% of original size

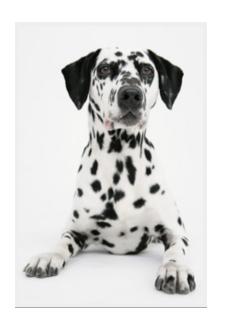
```
SET postgis.gdal_enabled_drivers TO 'PNG JPEG';
DROP TABLE IF EXISTS tmp_out ;

CREATE TABLE tmp out AS
SELECT lo from bytea(0, ST_AsPNG(ST_Resize(rast, 0.25, 0.25))) AS loid, filename FROM pics;

SELECT lo export(loid, '/tmp/' || file_name || '-25.png')
FROM tmp_out;

SELECT lo unlink(loid)
FROM tmp_out;
```

25% resized image



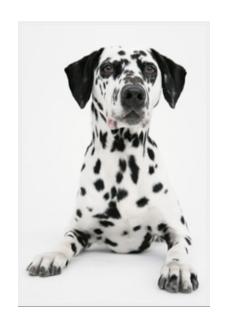


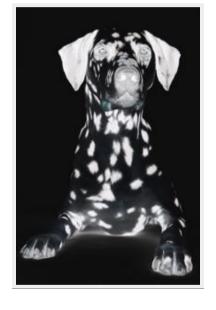
dalmation.jpg dalmation.jpg-25.png

CHANGE THE PIXEL BAND VALUES

A raster is an array of numbers. ST_Reclass lets you change the actual numbers by reclassifying them into ranges. This for example will allow you to reduce a 256 color image to 16 colors or change black spots to white spots.

DALMATION REVERSED





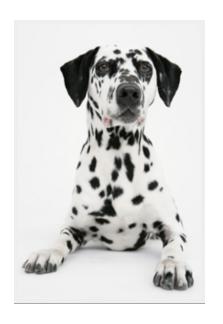
Before Reclass After Reclass

CROP THEM

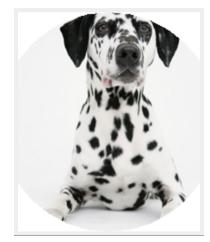
ST_Clip is the most commonly used function in PostGIS for raster. Here we buffer by 120 pixels from centroid of the picture and use that as our clipping region.

```
SELECT ST Clip(rast,
    ST Buffer(ST Centroid(rast::geometry), 120),
    '{0,0,0}'::integer[])
FROM pics
WHERE file_name = 'dalmatian.jpg';
```

DALMATION CROPPED

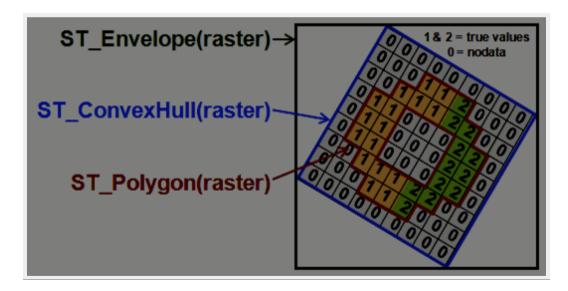


Before Crop After Crop



RASTER: ANALYZE ENVIRONMENTAL DATA

- Elevation
- Soil
- Weather



MIN, MAX, MEAN ELEVATION ALONG A ROAD

There are several stats functions available for raster. You'll almost always want to use these in conjunction with ST_Clip and ST_Count.

```
WITH estats AS
 (SELECT sld name, ST Count(clip) AS num pixels, ST SummaryStats(clip) AS ss
FROM
    nj ned AS e INNER JOIN
     (SELECT sld name, geom
         FROM nj roads
         WHERE std name IN( 'I-78', 'I-78 EXPRESS') ) AS r
             ON ST Intersects (geom, rast)
    , ST Clip (e.rast, r.geom) AS clip
SELECT sld name, MIN((ss).min) As min, MAX((ss).max) As max,
    SUM((s\overline{s}).mean*num pixels)/SUM(num pixels) AS mean
FROM estats
GROUP BY estats.sld name;
  sld_name | min
I-78 EXPRESS | -0.877017498016357 | 97.2313003540039 | 30.5347544396378
(2 rows)
Time: 422.456 ms
```

MANAGE DISCONTINUOUS DATE TIME RANGES WITH POSTGIS

A linestring can be used to represent a continous time range (using just X axis). A multi-linestring can be used to represent a related list of discontinous time ranges. PostGIS has hundreds of functions to work with linestrings and multilinestrings.

HELPER FUNCTION FOR CASTING LINESTRING TO DATE RANGES

```
CREATE FUNCTION to daterange (x geometry)
RETURNS daterange AS
$$

DECLARE
    y daterange;
    x1 date;
    x2 date;

BEGIN
    x1 = CASE WHEN ST_X(ST_StartPoint(x)) = 2415021 THEN '-infinity' ELSE 'J' || ST_X(ST_StartPoint(x)) = 2488070 THEN 'infinity' ELSE 'J' || ST_X(ST_ENDPoint(x)) EN    y = daterange(x1, x2, T[)');
    RETURN y;

END;
$$

LANGUAGE plpgsql IMMUTABLE;
```

HELPER FUNCTION FOR CASTING DATE RANGE TO LINESTRING

```
CREATE FUNCTION to linestring (x daterange)
RETURNS geometry AS
$$

DECLARE
    y geometry(linestring);
    x1 bigint;
    x2 bigint;

BEGIN

    x1 = to char(CASE WHEN lower(x) = '-infinity' THEN '1900-1-1' ELSE lower(x) F
    x2 = to char(CASE WHEN upper(x) = 'infinity' THEN '2100-1-1' ELSE upper(x) EN
    y = ST GeomFromText('LINESTRING(' || x1 || ' 0, ' || x2 || ' 0)');
    RETURN y;

END;
$$
LANGUAGE plpgsql IMMUTABLE;
```

COLLAPSING OVERLAPPING DATE RANGES

Result is single linestring which maps to date range

COLLAPSING DISCONTINUOUS / OVERLAPPING RANGES

Result is a multi-linestring which we dump out to get individual date ranges

COLLAPSING CONTIGUOUS DATE RANGES

Result is a linestring which we dump out to get individual date range

CREATE AN AGGREGATE FUNCTION

SUPER COLLAPSE

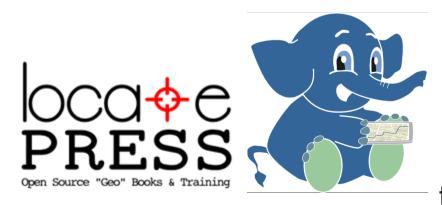
```
WITH z (id,x,grade) AS (
    VALUES
        ('alex', to linestring(daterange('2017-1-2','2017-1-3','[)')),'A'),
        ('alex', to linestring (daterange ('2017-1-1', '2017-1-6', '[)')), 'B'),
        ('alex', to linestring (daterange ('2017-1-5', '2017-1-8', '[)')), 'C'),
        ('alex', to linestring (daterange ('2017-1-1', '2017-1-9', '[)')), 'X'),
        ('beth', to linestring (daterange ('2017-1-1', '2017-1-3', '[)')), 'A'),
        ('beth', to linestring (daterange ('2017-1-5', '2017-1-9', '[)')), 'B'),
        ('beth', to linestring (daterange ('2017-1-1', '2017-1-9', '[)')), 'X')
SELECT
    a.id,
    to daterange (a.u) AS period,
    MIN (b.grade) AS grade
FROM
    (SELECT id, (ST Dump(ST Union(x))).geom AS u FROM z GROUP BY id) a
    INNER JOIN
    z b
    ON a.id = b.id AND ST Intersects(a.u,b.x) AND NOT ST Touches(a.u,b.x)
GROUP BY a.id, a.u
id
      | period
                                 | grade
alex | [2017-01-01,2017-01-02) | B
alex | [2017-01-02,2017-01-03) | A
alex | [2017-01-03,2017-01-05)
alex | [2017-01-05,2017-01-06]
alex | [2017-01-06,2017-01-08]
alex | [2017-01-08,2017-01-09)
beth | [2017-01-01,2017-01-03]
beth | [2017-01-03,2017-01-05)
beth | [2017-01-05,2017-01-09) | B
(9 rows)
```

```
WITH
    z (id,x,grade) AS (
         VALUES
         ('alex', to linestring(daterange('2017-1-2','2017-1-3','[)')),'A'),
         ('alex', to linestring (daterange ('2017-1-1', '2017-1-6', '[)')), 'B'),
         ('alex', to linestring (daterange ('2017-1-5', '2017-1-8', '[)')), 'C'),
         ('alex', to linestring (daterange ('2017-1-1', '2017-1-9', '[)')), 'X'),
         ('beth', to linestring (daterange ('2017-1-1', '2017-1-3', '[)')), 'A'),
         ('beth', to linestring (daterange ('2017-1-5', '2017-1-9', '[)')), 'B'),
         ('beth', to linestring (daterange ('2017-1-1', '2017-1-9', '[)')), 'X')
    ),
    w AS (
         SELECT
             a.id.
             a.u AS x,
             MIN(b.grade) AS grade
         FROM
             (SELECT id, (ST Dump(ST Union(x))).geom AS u FROM z GROUP BY id) a
             INNER JOIN
             z b
             ON a.id = b.id AND ST Intersects(a.u,b.x) AND NOT ST Touches(a.u,b.x)
         GROUP BY a.id, a.u
SELECT
    id,
id
       period
                                 I grade
alex I
       [2017-01-01,2017-01-02)
       [2017-01-02,2017-01-03)
alex |
alex |
       [2017-01-03,2017-01-06]
                                 l B
alex |
       [2017-01-06, 2017-01-08)
                                 I C
       [2017-01-08,2017-01-09)
alex |
       [2017-01-01,2017-01-03)
beth |
       [2017-01-03,2017-01-05)
beth |
       [2017-01-05,2017-01-09] | B
beth |
8 rows)
```

ROUTING WITH PGROUTING

Finding least costly route along constrained paths like roads, airline routes, driving distance analysis, fleet routing based on time constraints, and many more.

Buy our upcoming book (Feature complete Preview Edition available) pgRouting: A Practical Guide http://locatepress.com/pgrouting



to find out more

TRAVELING SALES PERSON PROBLEM

Routing nuclear power plant inspector

```
WITH
T AS (SELECT *
FROM pgr eucledianTSP($$SELECT id, ST X(geom) AS x, ST_Y(geom) AS y
FROM nuclear_power_plants$$, 19, 19)

SELECT T.seq, T.node AS id, N.name, N.geom, N.country
FROM T INNER JOIN nuclear_power_plants N ON T.node = N.id

ORDER BY seq;

seq | id | name | country

1 | 19 | Dukovany Nuclear Power Station | Czech Republic
2 | 20 | Temelin Nuclear Power Station | Czech Republic
3 | 45 | Isar Nuclear Power Plant | Germany
4 | 44 | Gundremmingen Nuclear Power Plant | Germany
5 | 46 | Neckarwestheim Nuclear Power Plant | Germany
6 | 47 | Philippsburg Nuclear Power Plant | Germany
150 | 55 | Higashid+iri Nuclear Power Plant | Germany
151 | 64 | Tomari Nuclear Power Plant | Japan
152 | 19 | Dukovany Nuclear Power Station | Czech Republic

The Action of the Action | Czech Republic

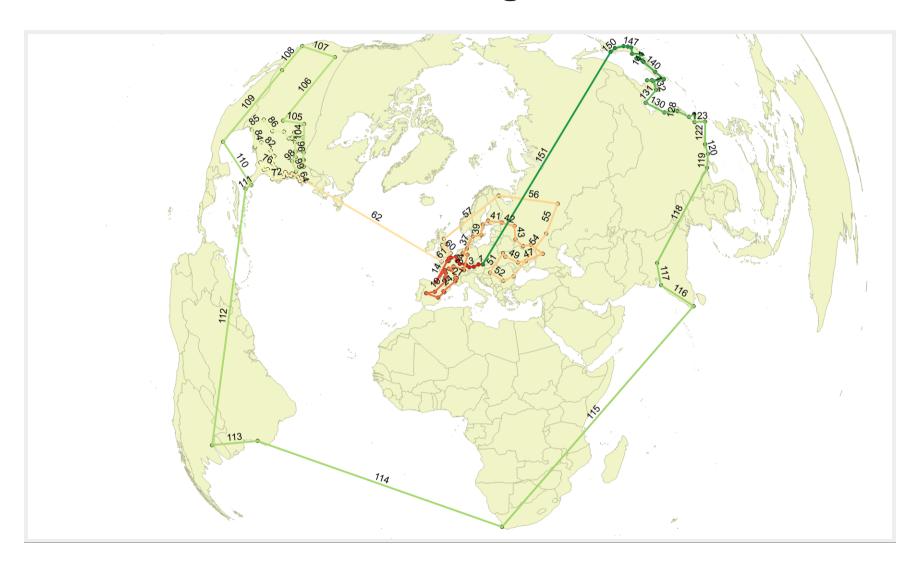
Czech Republic

Czech Republic

Czech Republic
```

Time: 48.706 ms

TSP IN QGIS



CATCHMENT AREAS: DRIVE TIME DISTANCE

What areas can a fire station service based on 5 minute drive time.

ALPHASHAPE AREA OUTPUT IN QGIS

Overlaid on roads network and with fire station location starred



DIJKSTRA: FINDING OPTIMAL ROUTE

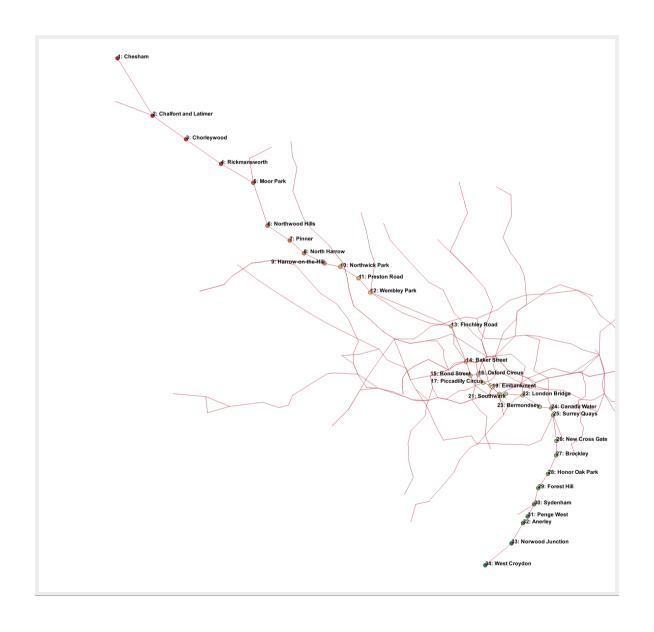
Fastest path from Chesham to West Croydon

Time: 36.509 ms

```
SELECT seg, S.station, L.name, round((cost * .000621371)::numeric,2) AS miles
FROM
     pgr dijkstra('
          SELECT gid AS id, source, target, length AS cost
               FROM london tube lines',
           (SELECT station id FROM london stations WHERE station = 'Chesham'), (SELECT station id FROM london stations WHERE station = 'West Croydon'),
          false
     ) R
     INNER JOIN london stations S ON R.node = S.station id
     LEFT JOIN london Tube lines L ON R.edge = L.gid
ORDER BY R.seq;
  1 | Chesham | Metropolitan | 3.38
  2 | Chalfont and Latimer | Metropolitan | 2.07
  3 | Chorleywood | Metropolitan | 2.15
 12 | Wembley Park | Metropolitan | 4.42
13 | Finchley Road | Metropolitan | 1.93
14 | Baker Street | Jubilee | 0.72
 19 | Embankment | Bakerloo | 0.49
 20 | Waterloo | Jubilee | 0.33
21 | Southwark | Jubilee | 0.86
 22 | London Bridge | Jubilee
                                          1.04
 23 | Bermondsey | Jubilee | 0.58
                          | Overground | 0.37
 24 | Canada Water
 31 | Penge West | Overground | 0.43
32 | Anerley | Overground | 1.16
33 | Norwood Junction | Overground | 1.72
 34 | West Croydon | NULL | 0.00
(34 rows)
```

DIJKSTRA: FINDING OPTIMAL ROUTE

London Tubes optimal path



LINKS OF INTEREST

- PostGIS
- Planet PostGIS
- pgRouting
- PostGIS.US our book site, includes code and data used in our books.

THE END

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