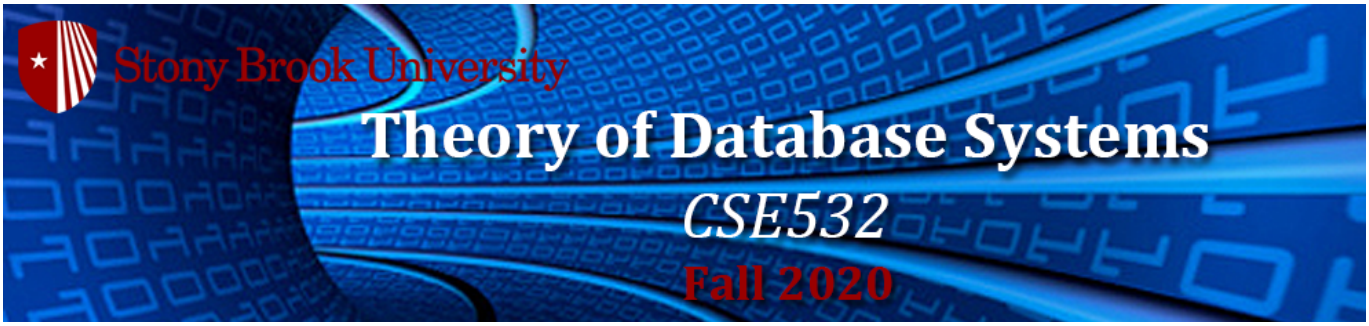

Home

Home
Syllabus
Schedule
Homework
Textbooks
Examples
Instructor
Resources

Homework >

Homework3: Spatial Queries

References:

- Please go through DB2 Spatial Tutorial (I, II) before you begin this homework.
- Please also read spatial query examples.
- DB2 Spatial Online Reference (V11.5) PDF Version

We will have two datasets for the homework.

The first dataset is from TIGER dataset provided by US Census Bureau. We will download the US counties in shapefile format. A local copy of the shapefile is also available here.
Once you unzip the file, you can run the following command under db2 command line to find metadata for the format. (Documentation of the dataset can be found here.)
db2se shape_info -fileName tl_2015_us_county.shp

The shapefile has a multipolygon object to represent the boundary of each county.

Dataset 2. COVID-19 testing sites in Long Island (Suffolk County and Nassau County) is available at COVID19testingsites.csv, including the name, address, zip code, and geolocation (latitude, longitude).

```
Name, address1, address2, city, state, zipcode, country, latitude, longitude
```

Task 1. (4 points) **Setup the database.**

a. Enable the sample database (or your own database) for spatial support:

```
db2se enable_db sample
```

b. Load the counties dataset using the import SQL file: *db2 -tf import_counties.sql*

c. Create two tables for covid-19 testing sites. The first one will have all columns in the csv file, without a spatial column:

```
cse532.siteorginal(Name, address1, address2, city, state, zipcode, country, latitude, longitude)
```

The second table cse532.site will replace columns (latitude, longitude) with a spatial column geolocation DB2GSE.ST_POINT:

```
cse532.site(Name, address1, address2, city, state, zipcode, country, geolocation)
```

Please put both CREATE TABLE statements plus the following statement (registering the spatial column of site table in the spatial database) in createsitetables.sql:

```
!db2se register_spatial_column sample
-tableSchema      cse532
-tableName        site
-columnName       geolocation
-srsName          nad83_srs_1
;
```

Run the codes in the file to create the tables:

```
db2 -tf createsitetables.sql
```

d. Write a data loading script siteload.sql to load data from the csv file into cse532.siteoriginal table.

~~db2 load from "C:\myfolder\Health_Facility_General_Information.csv" of del MESSAGES load.msg INSERT INTO cse532.facilityoriginal~~

e. Write a SQL script siteinsert.sql to insert data into cse532.site~~original~~ by selecting data from cse532.siteoriginal table and converting (*Latitude, Longitude*) attributes into DB2GSE.ST_POINT type with srs_id 1 for *geolocation* attribute in cse532.site.

f. Update the createindexes.sql to add additional indexes besides spatial indexes that may accelerate the queries below.

Task 2. (3 points) **Write a query nearestsite.sql to find three closest COVID-19 sites from "2799 Horseblock Road Medford, NY 11763"(40.824369, -72.993983) (latitude, longitude)**. Please return location and distance in your result. You can use unit 'KILOMETER', 'METER', or 'STATUTE MILE' for distance measurement.

Nearest neighbor search is not directed supported by DB2. You can use ST_BUFFER to create a buffered area (polygon/circle) from a point within a certain distance and search only stores within the buffer. Note that 0.25 degree is roughly 10 miles. For all the datasets, we use spatial reference nad83_srs_1 with srs ID as 1.

You can find information here on functions such as ST_POINT, ST_BUFFER, ST_WITHIN or ST_CONTAINS, and ST_DISTANCE.

3. (3 points). Write a query suffolksitecount.sql to return the total number of sites in Suffolk County of New York State.

Submission.

Please zip your SQL scripts, results, and a README file. A result file can be a text file or a screenshot.

Please go to blackboard, and submit it under homework 3.

Subpages (1): querytime

COVID19testingsites.csv (9k)	Fusheng Wang, Oct 15, 2020, 10:08 PM	v.2
createindexes.sql (0k)	Fusheng Wang, Oct 18, 2020, 7:49 PM	v.1
import_counties.sql (1k)	Fusheng Wang, Oct 18, 2020, 7:33 PM	v.1