Spatial Database Management GEP 664 / GEP 380 Class #1: Introduction

Frank Donnelly

Dept of EEGS, Lehman College CUNY

Spring 2019

Today's Topics

Course Overview

The Syllabus

Technical Details

Next Class

Spatial Databases

PostgreSQL & PostGIS

Database that has been optimized to store spatial data and perform geographic functions.

Organization and storage: keep related geographic features and data tables together

Relational database: utilizes all the benefits of this structure

Multi-user: many people can read and write simultaneously

Automation and validation of tasks: connect to data with many tools, control over data quality

Work with big datasets: when desktop software and online tools fall flat

▲□▶▲圖▶▲臣▶▲臣▶ 臣 の�?





▲□▶▲□▶▲≡▶▲≡▶ Ξ のQで

Data Structure

An organized collection of data designed for efficiently answering questions and storing information.

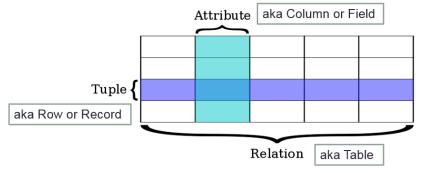


Image source: https://en.wikipedia.org/wiki/Relational_database

Values for specific attributes are drawn from an allowable set called a domain. Attributes are assigned data types, which limits the allowable values and operations that can be performed.

- Variable characters / Text (string)
- Integers (whole numbers)
- Reals / Floats (decimal numbers)
- ► Time and Date

・ロト・日本・モー・ 日 うらの

Keys and Joins

SQL Data Manipulation

This structure helps to insure the integrity of the data and makes it possible to relate values in one table to values in another using an attribute they hold in common: a unique ID code called a primary key.

	FID	Shape *	STATE	COUNTY	NAME	LSAD	^	FIPS	County	TotalDeathRate	Heart	Neoplasms	Ī
	32	Polygon	36	001	Albany	06		001	Albany	938.2	315	211.8	1
	41	Polygon	36	003	Allegany	06		003	Allegany	897.7	263	203.7	1
	58	Polygon	36	059	Nassau	06		005	Bronx	702.9	251	148.1	1
	45	Polygon	36	007	Broome	06		007	Broome	1048.1	303	231.2	
Γ	40	Polygon	36	009	Cattaraugus	06		009	Cattaraugus	1089.2	413.6	217.6	1
	12	Polygon	36	011	Cayuga	06		011	Cayuga	854.5	278.3	189.2	
	38	Polygon	36	013	Chautauqua	06		013	Chautauqua	1038.9	328.6	229.5	1
	47	Polygon	36	015	Chemung	06		015	Chemung	1001.3	295.6	238.9	1
Γ	35	Polygon	36	017	Chenango	06		017	Chenango	1060.5	441.6	223.7	1
	1	Polygon	36	019	Clinton	06	~	019	Clinton	763.4	204	178.3	
						>		<				>	

SQL is the language for creating and manipulating relational databases; originally based on relational algebra, it uses declarative commands in English.

SELECT county_name, pop AS population FROM countypop WHERE state='NY' AND pop > 50000 ORDER BY pop;

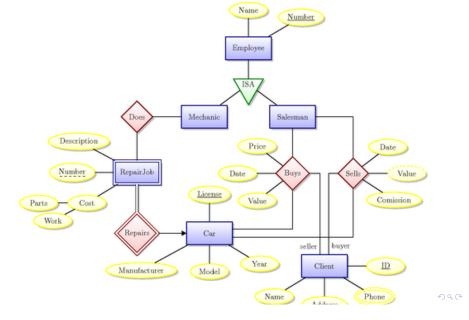
▲□▶▲圖▶▲臣▶▲臣▶ 臣 の�?

SQL Data Design

Database Design

The structure of the relational database and the SQL language were designed to be independent of any specific hardware or software.

CREATE TABLE countypop (geoid varchar(5) CHECK (char_length(geoid) = 5), state varchar(2), county_name text, pop integer, CONSTRAINT gidkey PRIMARY KEY (geoid));



Data Processing i.e. Extract, Transform, Load

BP_2014_00CZ1_with_ann - Notepad		idit Data - Po	ostgreSQL 9.3 (localhost:5432) - gep664 - i	yszbp2014				- C - X
File Edit Format View Help	FILE	Edit Viev	w Iools Help					
GEO. id, GEO. id2, GEO. display-label, NAICS. id, NAICS. display- label, YEAR, id, ESTAB, EMP, PAYOTRI, PAYANN		1 🤊 🔿 🛯 🖻	🛯 🚳 🐨 🍞 👔 100 rows 👻					
Geographic identifier code.Id2.Geographic area name.2012 NAICS code.Meaning of 2012		zipcode	zipname		employees integer	empflag character v	payrol_tho	payflag character v
NAICS code, Year, Number of establishments, Paid employees for pay period including		[PK] chai 00501	ZIP 00501 (Moltaville, NY)	integer	integer	character v	nteger	character v
March 12 (number), "First-quarter payrol] (\$1,000)", "Annual payrol] (\$1,000)"					197	D	13178	5
86100000500501,00501,"ZIP 00501 (Holtsville, NY)",00,Total for all		06390	SIP 06390 (Fishers Island, N					
sectors,2014,4,b,D,D 8610000US06390,06390,"ZIP 06390 (Fishers Island, NY)".00.Total for all		10001	SIP 10001 (New York, NY)	7369	150381		8907912	
sectors, 2014, 40, 197, 246, 13178			SIP 10002 (New York, NY)	2972	21459		671500	
8610000US10001,10001,"ZIP 10001 (New York, NY)",00,Total for all	5	10003	SIP 10003 (New York, NY)	4333	109860		6887483	
sectors, 2014, 7369, 150381, 2207824, 8907912		10004	SIP 10004 (New York, NY)	1554	62016		11733524	
86100000\$10002,10002,"ZIP 10002 (New York, NY)".00.Total for all		10005	SIP 10005 (New York, NY)	1321	42344		6761564	
sectors, 2014, 2972, 21459, 155145, 671500		10006	SIP 10006 (New York, NY)	855	19726		2300230	
8610000ÚS100Ó3,10Ó03,"ZÍP 10003 (New York, NY)",00,Total for all	5	10007	SIP 10007 (New York, NY)	1451	28772		2798122	
sectors, 2014, 4333, 109860, 1754943, 6887483		10008	SIP 1000B (New York, NY)	15			17464	
86100000\$10004,10004,"ZIP 10004 (New York, NY)",00,Total for all		10009	SIP 10009 (New York, NY)	1248	9600		411148	
sectors, 2014, 1554, 62016, 5584559, 11733524		2 10010	SIP 10010 (New York, NY)	3137	74039		7860140	
86100000510005,10005,"ZIP 10005 (New York, NY)",00,Total for all		3 10011	SIP 10011 (New York, NY)	3958	63474		4164642	
sectors,2014,1321,42344,2708958,6761564 8610000US10006,10006,"ZIP 10006 (New York, NY)",00,Total for all				2983				
sectors 2014,855,19726,645737,2300230		10012	ZIP 10012 (New York, NY)		41920		2444313	
8610000US10007,10007,"ZIP 10007 (New York, NY)".00.Total for all		5 10013	SIP 10013 (New York, NY)	5417	76461		7492878	
sectors.2014.1451.2872.894601.2798122		5 10014	SIP 10014 (New York, NY)	2448	43683		2774296	
8610000US10008,10008,"ZIP 10008 (New York, NY)".00.Total for all	1	10016	SIP 10016 (New York, NY)	5968	113168		9202349	
sectors .2014.15.f.4212.17464	1	3 10017	SIP 10017 (New York, NY)	4963	151614		19910777	
86100000\$10009,10009,"ZIP 10009 (New York, NY)".00.Total for all	1	10018	SIP 10018 (New York, NY)	6808	136779		10547799	
sectors 2014 1248 9600 93619 411148			man 40040 (m m) mm)	4000	4 4 4 4 0 7		00000404	

Spatial Databases

Geographic vector features stored as series of coordinates in a geometry data type. Geometry sub-types:

- ▶ POINT(0 0)
- LINESTRING(0 0,1 1,1 2)
- POLYGON((0 0,4 0,4 4,0 4,0 0),(1 1, 2 1, 2 2, 1 2,1 1))
- MULTIPOINT((0 0),(1 2))
- MULTILINESTRING((0 0,1 1,1 2),(2 3,3 2,5 4))
- MULTIPOLYGON(((0 0,4 0,4 4,0 4,0 0),(1 1,2 1,2 2,1 2,1 1)), ((-1 -1,-1 -2,-2 -2,-2 -1,-1 -1)))
- GEOMETRYCOLLECTION(POINT(2 3),LINESTRING(2 3,3 4))

Image source: http://www.texample.net/tikz/examples/entity-relationship-diagram/

▲□▶ ▲□▶ ▲臣▶ ▲臣▶ 三臣 - のへで

Spatial Relationships & Analysis

FROM boroughs, stations

WHERE bname='Bronx' AND

SELECT bname, stop_name, trains

ST_Within (stations.geom, boroughs.geom);

When features have geometry they can be compared spatialy.

Geometry columns are referenced with internal metadata tables that store coordinate systems.

Table Relationships

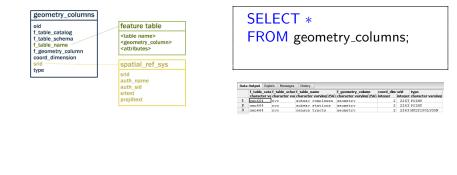


 Image: Contains
 Image: Contains

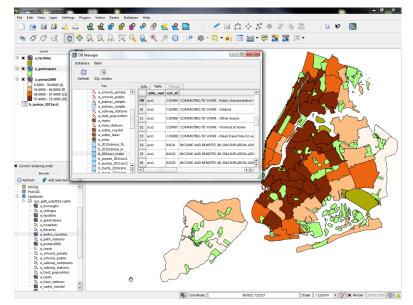
Image source: https://www.gaia-gis.it/spatialite-2.1/SpatiaLite-manual_html 🛓 🐌 👔 🔊 Q (>

Proximity Analysis

Adjacency, Buffers, Distance, Distance Within, Nearest Neighbor. Spatial relationships can be evaluated quickly using spatial indexes.



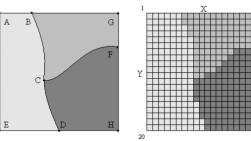
Spatial Data Organization & Desktop GIS



▲□▶▲□▶▲臣▶▲臣▶ ■ のへの

▲□▶▲圖▶▲臣▶▲臣▶ 臣 の�?

- Roles and permissions for users
- Triggers and transactions for processes
- ► Fine tuning and performance
- Backup and restoration



Vector image

Raster image

ector			Raster	
olygon ID	Coordinates	Soil Type	Grid Ref.	ltem
	A,B,C,D,E	Chalk	x=1, y=1	Chalk
	B,C,F,G	Clay	X=2, y=1	Chalk
	C,F,H,D	Gravel	X=3, y= 1	Chalk
			X=4 etc.	
			X=20, y=20	Gravel

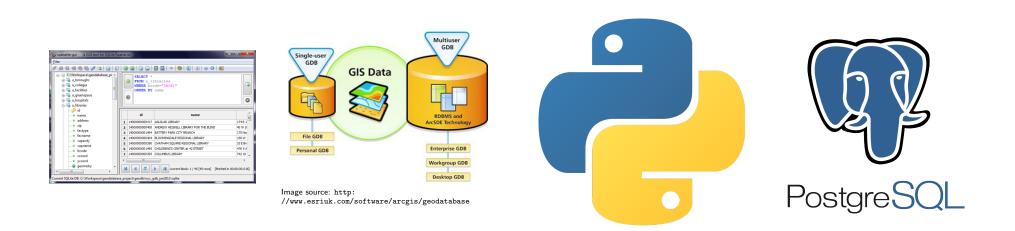
Image from http://www.arts-humanities.net/wiki/gis_geographic_information_system_archaeology

▲□▶ ▲□▶ ▲目▶ ▲目▶ 目 のへの

・ロト・日本・日本・日本・日本・日本

Other Database Formats Spatialite for desktop, ArcGIS formats

Programming / Scripting Languages Python and PostgreSQL



Classes 1 through 5 Essentials of Relational Databases (Jan - Feb)

Classes 6 through 10 Essentials of Spatial Databases (Mar - Apr)

Spring Break Apr 23

Classes 11 through 13 Special Topics (Apr - May)

Class 14 and Final Final Projects (May)

For each class:

- 1. Hand in assignments
- 2. Lecture and discussion
- 3. Exercises
- 4. Break

Today's Topics

- 5. Lecture and discussion
- 6. Exercises
- 7. Return previous assignments and discuss
- 8. Lab time to begin next assignments

My Background Geography & Library & Information Science

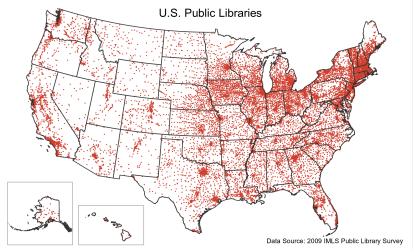


Image source: https://atcoordinates.info/tag/public-libraries/

Course Overview

The Syllabus

Technical Details

Next Class

The Syllabus	Today's Topics	
	Course Overview	
Course website: https://spatdb.commons.gc.cuny.edu/	The Syllabus	
	Technical Details	
	Next Class	
	۹. (۲.	< □ > < □ > < 芝 > < 芝 > < 芝 > < ジ < ぐ
	Auchitestum	

Interfaces

Architecture

Databases exist as discrete objects outside of a specific interface. You can interact with a CLI or GUI.



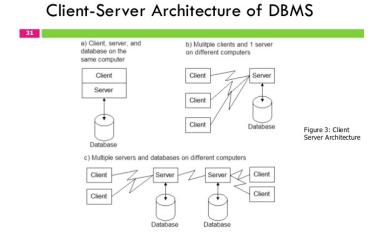


Image source: http://www.slideshare.net/b15ku7/chapter-1-fundamentals-of-database-management-system

Today's Topics	Due Next Class
Course Overview	The following are due at the beginning of our next class:
The Syllabus	Assignment #1 Posted on the course website (under Assignments)
	Readings for Class #2 Listed in the syllabus, in the <i>Practical SQL</i> book
Technical Details	Note: There is overlap in course content and readings for
Next Class	classes 2 & 3

< □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □

・ロ・・母・・ヨ・・ヨー ひゃぐ