GEOG 28702

Geographic Information Science II

Winter 2020

Schedule and Location:
Tu Th, 2:00-3:20pm
Rosenwald Hall

Instructor:
Marynia Kolak, mkolak@uchicago.edu,
Office location: 232 Searle Lab
Office hours: M 2:45-4:15 pm & by appointment

Overview

This course investigates the theory and practice of infrastructure and computational approaches in spatial analysis and GIScience. Geocomputation is introduced as a multidisciplinary systems paradigm necessary for solving complex spatial problems and facilitating new understandings. Students will learn about the elements of spatial algorithms and data structures, geospatial topologies, spatial data queries, and the basics of geodatabase architecture and design.

In this methods course, each class builds on topics discussed and practiced from the previous one. Students will learn and practice *SQL* and the *R* programming language to implement several techniques learned. It is thus critical that you attend all classes, complete all required assignments and quizzes, and participate in the discussion (both in-class and in online forums). All materials for the course will be made available on the course *Canvas* site. This site will also be the main means of communication.

This course maintains a no-paper policy (meaning no paper products), so all assignments etc. should be delivered in digital form (as a pdf – no Word documents!) to the *Canvas* site.

Course Requirements

Each class session will have a combination of activities. **Personal computers (laptops) must be brought to every class for in-class coding and related exercises.** Each topic will be introduced with a lecture, synthesized with a problem-based group exercise, and further examined with an individual lab practical. In addition to active participation in class sessions, students are expected to complete assignments, quizzes, and examinations. Assignments are short papers (2-pages) that require an overview of work completed, carefully labeled diagrams and figures, and thoughtful discussion on process. Quizzes will review key concepts in the class material covered. Lab practicals require confirmation of work completed. A final project will cover materials presented throughout the course, and will serve as three submissions combined.

Grading

Final grades are assigned by composition of work earned in participation, quizzes, assignments, labs, and class project. Unless otherwise noted, each assigned work is graded on a 0-100 point scale, and then weighted according to percentage worth for the final grade. Weighted grades as they are in current standing will be posted "live" on *Canvas* throughout the course for easy tracking.

Final Grade Composition:

Discussions 15%
Assignments 15%
Labs 15 %
Quizzes 15%

Project 40%
Proposal 10%
Data Model 10%
Lightning Talk 5%
Final Report 15%

Participation.

Attendance and thoughtful participation are expected of all students. Both in-person and online participation is expected. Total worth is 15%.

Quiz.

Three quizzes will be assigned on Canvas to review materials from the course. Each quiz is worth 5% for a total of 15%. No late submissions are allowed, extenuating circumstances permitting.

Assignments.

There are 3 assignments in total. Each is worth 5%, for a total of 15%. For every day an assignment is late, there will be a 5pt deduction.

Labs.

There are six labs. Each is worth 3%, for a total of 15% (with lowest lab grade dropped). No late submissions are allowed, extenuating circumstances permitting.

Class Project.

The class project will involve the development of a spatial database to model an urban or rural environment. The proposal (2-pages) must describe and defend datasets chosen to model the environment and define the purpose of the project, and is worth 10%. The data model must include a full entity relationship data model showing data structures and linkages of all data attributes used, and is worth 10%. A lightning talk in final week of class should provide a 5-minute overview of the work. The final report must include a solution framework with documented workflow diagrams and description of process, and is worth 15%. The total project is worth 40%.

The Grading Schema is the following standard: A = 90-100%, B = 80-89%, C = 70-79%, D = 60-69%.

Software

The class uses only open source software (free and cross-platform). You are required to install it on **your own machine.** Everything can be readily downloaded from the web. Additional software may be added to the course, but will be done so with advanced notice. Please contact the instructor if you cannot access your own machine (desktop or laptop) for this course.

See also the Lab 1-2 installation overview (available on *Canvas)* for more details.

Books

PostGis in Action, 2nd edition, Hsu and Obe (2015)

Additional required readings will be made available and posted on the *Canvas* course page, including all lecture material, class slides, lab practicals, and external links.

Tentative Course Outline

May be subject to change.

Section 1 Spatial Database Foundations

Week 1. Spatial Data in the Geocomputational Paradigm
- (Jan 7 M) Defining the Geocomputational Paradigm, Course Overview
- (Jan 9 W) Lab 1. Open Spatial Infrastructure Initialization

Readings:
- Openshaw, S., & Abrahart, R. J. (1996). Geocomputation. *In the proceedings of The 1st International Conference on GeoComputation*, University of Leeds, Leeds, England (Vol. 1, pp. 665-666).
- Couclelis, H. (1998, November). Geocomputation in context. *In Geocomputation: A primer* (pp. 17-29). John Wiley Chichester.
- Harris, R., O'Sullivan, D., Gahegan, M., Charlton, M., Comber, L., Longley, P., ... & Arribas-Bel, D. (2017). More bark than bytes? Reflections on 21+ years of geocomputation. *Environment and Planning B: Urban Analytics and City Science*, 44(4), 598-617.

Due Friday: Lab 1

Week 2. The Basics of a Spatial Database
- (Jan 14 M) Introduction to Spatial (and Non-Spatial) Databases, Spatial Data Types Review
- (Jan 16 W) Spatial Reference Systems, *Practice:* Developing your First Database

Readings:
- Beaulieu , A. (2009) "Chapter 1: A Little Background." *Learning SQL, 2nd Edition.* O'Reilly Media.
- Hsu, L. and Obe, R. (2015). "Chapter 1: What is a Spatial Database?" *PostGIS in Action.* Manning.

Due Friday: Quiz 1

Week 3. Database Data Types, Design, and Diagramming
- (Jan 21 M) Introduction to SQL, Database Model Type and Design, *Practice:* ER Diagrams
- (Jan 23 W) Lab 2. Simple SQL Queries in a Spatial Database

Readings:
- "Chapter 2. The SQL Language." pSQL Documentation.
- "What is an ER Diagram?" LucidChart.
- Brady, M., & Loonam, J. (2010). Exploring the use of entity-relationship diagramming as a technique to support grounded theory inquiry. *Qualitative Research in Organizations and Management: An International Journal.*

Due Friday: Lab 2

Section 2 Spatial Data Modeling Frameworks

Week 4. Modeling Environments Using Data

- (Jan 28) Data Representations in GIS, *Practice:* Entity Relationship Diagrams
- (Jan 30) Field and Object View Spatial Frameworks, *Practice:* Spatial Views

Readings:

- Couclelis, H. (1992). People manipulate objects (but cultivate fields): beyond the raster-vector debate in GIS. In *Theories and methods of spatio-temporal reasoning in geographic space* (pp. 65-77). Springer, Berlin, Heidelberg.
- Goodchild, M. F., Yuan, M., & Cova, T. J. (2007). Towards a general theory of geographic representation in GIS. *International journal of geographical information science*, 21(3), 239-260.

Due Friday: Project Proposal, Assignment 1

Week 5. Systems Thinking in Spatial Database Design

- (Feb 4) Spatial DB Extract, Transform, and Load (ETL) Workflows, *Practice:* Concept Mapping
- (Feb 6) Lab 3. Spatial Database Generation using Automated Data Processing

Due Friday: Quiz 2. Spatial and Non-Spatial Data Queries, Lab 3

Readings:

- Cohen, T. L., & Baitty, J. R. (2005, June). Health Resources and Services Administration Geospatial Data Warehouse. Integrating Spatial and Tabular Extract, Transform, and Load processes. In *ESRI 2005 International Users Conference* July 24-29, 2005.
- McGuire, M., Gangopadhyay, A., Komlodi, A., & Swan, C. (2008). A user-centered design for a spatial data warehouse for data exploration in environmental research. *Ecological Informatics*, 3(4-5), 273-285.

Week 6. Formalizing Spatial Data Infrastructures

- (Feb 18) Spatial Data Infrastructures, *Practice:* ETL workflows with Open data
- (Feb 20) Lab 4. Basics of Buffer & Location Analysis

Readings:

- Hu, Y. and Li, W. (2017). Spatial Data Infrastructures. The Geographic Information Science & Technology Body of Knowledge (2nd Quarter 2017 Edition), John P. Wilson (ed.). DOI: 10.22224/gistbok/2017.2.1
- Budhathoki, N. R., & Nedovic-Budic, Z. (2008). Reconceptualizing the role of the user of spatial data infrastructure. GeoJournal, 72(3-4), 149-160.

Due Friday: Assignment 2. Server-Side Spatial Data Calculations, Lab 4

Section 3 Geoprocessing and Geoanalytics with Spatial DBs

Week 7. **Defining Complex Spatial Data Relationships**

- (Feb 11) Tree Data Structures and Spatial Topologies *Practice*: Geometry Operations in POSTGIS
- (Feb 13) Spatial Search Algorithms, *Practice:* Spatial Relationships & Joins in POSTGIS

Readings:

- Chen, J., Li, C., Li, Z., & Gold, C. (2001). A Voronoi-based 9-intersection model for spatial relations. *International Journal of Geographical Information Science*, 15(3), 201-220.
- Agafonkin, V. (2017) “A dive into spatial search algorithms.” Mapbox Blog.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1), 269-271.

Due Friday: Project Data Model

Week 8. **Geometry and and Geometric Processing Optimizations**

- (Feb 25), Geoprocessing Review, Query Performance Tuning, *Practice:* Nearest-Neighbor Searching
- (Feb 27) Lab 5. Spatial and Geometric Operations in POSTGIS

Readings:

- Song, M., Li, W., Zhou, B., & Lei, T. (2016). Spatiotemporal data representation and its effect on the performance of spatial analysis in a cyberinfrastructure environment–A case study with raster zonal analysis. *Computers & Geosciences*, 87, 11-21.
- Bartoszewski, D., Piorkowski, A., & Lupa, M. (2019, May). The comparison of processing efficiency of spatial data for PostGIS and MongoDB databases. In *International Conference: Beyond Databases, Architectures and Structures* (pp. 291-302). Springer, Cham.

Due Friday: Quiz 6. Diagramming Data Workflows, Lab 5

Week 9. **Extending Spatial DBs as Stack Solutions**

- (Mar 4) Integrating PostGIS with QGIS, R, and more, *Practice:* Topologies for Routing
- (Mar 6) Lab 6. Service Catchment Areas with pgrouting, POSTGIS, and R

Readings:

- Olivier Bonin. How to make R, PostGIS and QGis cooperate for statistical modelling duties: a case study on hedonic regressions. *Open Source Geospatial Research and Education Symposium (OGRS)*, Oct 2012, Yverdon-les-Bains, Switzerland. pp.1. Ffhalshs-00737397f
- Choosumrong, S., Raghavan, V., & Bozon, N. (2012). Multi-criteria emergency route planning based on analytical hierarchy process and pgRouting. *Geoinformatics*, 23(4), 159-167.
- Basille, M., Urbano, F., & Conway, J. (2014). A step further in the integration of data management and analysis: Pl/R. In *Spatial Database for GPS Wildlife Tracking Data* (pp. 213-229). Springer, Cham.

Due Friday: Assignment 3. Mastering Data Workflows

Section 4 Tying it Together: Final Projects Due

Week 10 **Course Conclusions**

- (Mar 11) Project Lightning Talks
- (Mar 13) GeoComputation: Past, Present, Future

Readings:

- Rey, S. J. (2014). Open regional science. *The Annals of Regional Science*, 52(3), 825-837.

Due **Monday Mar 18**: Final Project