UP 494 Advanced Data Science for Planners

Instructor: Dr. Fang Fang (fangf@illinois.edu)

Weekly synchronous lab session: Friday 09:30 AM - 10:50 AM, Nevada lab

Lectures: asynchronous

Office Hours (online): Mon 9:00 AM - 10:00 AM Thursday 2:00 PM - 4:00 PM , Otherwise, by appointment

TA: Chaeyeon Han, email: ch55@illinois.edu **Office hours (online):** Tuesday 1:00-3:00PM

Course Overview:

Nowadays we are using new datasets and new technologies to experience, understand and plan cities. Data science is this multi-disciplinary field that involves these scientific methods, technologies, algorithms to extract and further understand knowledge or insights from all kinds of data. This is the advanced data science course for upper-level undergrads and graduate students in urban planning. No previous coding experiences are required. But students are expected to have the basic knowledge of planning theories, statistics, and GIS foundations. For the first half of the course, you will learn a set of basic concepts, skills, and tools in R for effective data processing. In addition, for the second half of the semester, more advanced and project-oriented topics will be covered toward social science e.g. spatial analysis, pattern analysis, and text analytics, with weekly readings and discussions. A final project report and presentation are expected. Our goal is not to become professional programmers, but students can use fundamental skills to understand and solve complex issues in the cities. This applied course contributes lots of technique skillsets for students, especially for their capstone/thesis/dissertation work.

This syllabus is subject to change by the instructor.

Course Outcomes:

- Prepare, collect, manipulate, query, and basic work with data at an intermediate level
- Perform data summarization, and statistical modeling
- Produce high-quality graphs, maps, and other types of deliveries as communication for planners
- Perform geospatial analysis in R
- Present and interpret results to foster reproducible research
- Parse and analyze text documents

Course Structure/Philosophy/Attendance

- I firmly believe that students learn via engagement and by doing. As a result, this will not be a pure lecture-based course. **It is important that you engage yourself during this class**. I will do my best to help you learn; however, it is imperative that you take ownership of your

education. You are responsible for all the works you did in this course. Come see me if you need help.

- This is a 16-week / full semester course. We have one lecture and one lab each week. **You should complete all exams, weekly discussions, assignments, and labs.**
- Each student is expected to devote 1-2 hours per week learning the asynchronous lectures and all types of contents, and 1-2 hours for labs exercise per week.
- Weekly contents will be posted every Monday by 9 AM.
- All the assignments, exams, discussions, and labs are mandatory.
- Online lab sections are delivered every Friday 9:30AM - 10:50AM using Zoom. The attendance of online live labs is highly recommended. Due to chaos of COVID-19 we do not check attendance especially we have students outside the U.S and in different time zones. Each student should attend lab sessions unless unavoidable conflict which must be notified to instructor ahead of time. I would also recommend you start working on each lab before Friday asap, so you can bring your questions during the online hours to troubleshoot. Students are expected to devote additional hours to finish the labs if needed.

Required Textbooks

- Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics. (2017) Editors: Thakuriah, Piyushimita (Vonu), Tilahun, Nebiyou, Zellner, Moira (Eds.), by Springer International Publishing.
- Beginning Data Science with R. (2014) Editor: Pathak, Manas A, by Springer International Publishing.
- R for Data Science import, tidy, transform, visualize and model data. By Hadley Wickham & Garrett Grolemund. Click here for a digital version.

Software

Students can install RStudio on their personal computers for free. Please refer to the lab0 instruction.

Lab Assignments and Late Work Policy:

In addition to the exams, you will be asked to complete 9 lab assignments. Assignments must be turned in via Compass submission. Unless otherwise stated, the lab assignments are due on 9:30 AM of the Friday that one week after they are assigned (e.g. a lab assigned on Aug. 28st will due on 9:30 AM of Sep. 4th). You should submit your assignment to Compass website. An assignment**, including lab assignments, mid-term exam, discussion, project proposal and final project**, submitted 24 hours or less after the due date will only be eligible for 80% of the maximum number of points allotted. Assignments submitted more than 24 hours but less than 48 hours after the due date will only be eligible for 60% of the maximum number of points allotted, and so on. Assignments submitted **more than 120 hours (or 5 days)** after the due date **will NOT be accepted and you will receive a zero on that assignment**. If you experience extenuating circumstances (e.g., you are hospitalized) that prohibit you from submitting your assignments on time, please let me know. I will evaluate these instances on a case-by-case basis. You are responsible to confirm each submission in Compass. **For any technical**

issues in Compass/Netid, you need to contact me in advance or email your assignment to me ASAP by the deadline. Otherwise, the late work policy will be strictly enforced.

Error/warning messages are very common in R, and these are **NOT** the valid excuses for late assignment submission. It is your responsibility to utilize resources (textbook, office hours, ask the instructor for help, online resources, etc.) to debug your code.

Exams:

You need to finish a take-home exam as the midterm, please check out the details below. **Make-up exams due to any other reason will be at the discretion of the instructor. You must notify me beforehand if you are going to miss an exam. I will not let you make up an exam without prior notification.**

Final project

All the student needs to finish a final project report with a presentation. The details will be posted on Compass later this semester.

Undergraduate and Graduate Students' Workload:

- For undergraduate students, the final project can be finished as a group. You should email the instructor about your group info by **11:59 PM, Oct 9th**.
- The graduate students must finish the project individually.

A project proposal is due by **09:30 AM, Oct 19th**. A final report is required as delivery by **09:30 AM, Dec 7th**. The details and requirements will be posted later in Compass.

Participation:

There is a discussion forum in Compass site where you need to post your thoughts each week based on the required reading assignments. Your timely posts (within 7 days) and replies are counted towards participation, which worth 100 points total. Let's try to be a vibrate community and hope these forums can be good resources for this course. Note simply or low-quality replies e.g. "I agree with Andrew" or "The article is very interesting" will **NOT** be accepted as participation.

Grade Point Distribution:

Assignments*9	500 Points Total
Mid-term take home Exam	200 Points
Final project	200 Points (20 proposal, 30 presentation, 150 report)
Participation (Online discussion)	100 Points (10 points/reading)
Total	**1000 Points**

Grade Scale:

Letter grade	Percentage	Points
A+	97–100%	>970
A	93–96.99%	>930
A–	90–92.99%	>900
B+	87–89.99%	>870
B	83–86.99%	>830
B–	80–82.99%	>800
C+	77–79.99%	>770
C	73–76.99%	>730
C-	70–72.99%	>700
D+	67–69.99%	>670
D	63–66.99%	>630
D-	60–62.99%	>600
F	0–59.99%	<600

Academic Integrity

We will follow Articles 1-401 through 1-406 of the Student Code. The provisions of the Student Code are applicable to this course. This rule defines infractions of academic integrity, which include but are not limited to cheating, fabrication, and plagiarism. You are responsible for following these guidelines. If you have any questions about whether something would be an infraction, consult with the instructor before proceeding.

Special Accommodations

We will accommodate students with documented disabilities. Please be familiar with the services and resources provided by Disability Resources and Educational Services (DRES) and visit (http://disability.illinois.edu/disability-resource-guide) for more information. Please inform the instructor of any requests at the beginning of the semester.

Feedback Response Time

I generally reply to email and discussion posts within 48 hours, except during holidays. I often reply much quicker, but you should not count on a same-day reply. Please plan accordingly so that you don't miss deadlines! I generally return assignments within one week after a discussion or assignment closes. If you would like to get help on an assignment ahead of the deadline, please email me as early as possible! I'm happy to give preliminary feedback and/or answer your questions.

Emergency Response Recommendations

Emergency response recommendations can be found at the following website: http://police.illinois.edu/emergency-preparedness/. I encourage you to review this website and the campus building floor plans website within the first 10 days of class. http://police.illinois.edu/emergency-preparedness/building-emergency-action-plans/.

Family Educational Rights and Privacy Act (FERPA)

Any student who has suppressed their directory information pursuant to Family Educational Rights and Privacy Act (FERPA) should self-identify to the instructor to ensure protection of the privacy of their attendance in this course. See https://registrar.illinois.edu/academic-records/ferpa/ for more information on FERPA.

Sexual Misconduct Policy and Reporting

The University of Illinois is committed to combating sexual misconduct. Faculty and staff members are required to report any instances of sexual misconduct to the University's Title IX and Disability Office. In turn, an individual with the Title IX and Disability Office will provide information about rights and options, including accommodations, support services, the campus disciplinary process, and law enforcement options.

Tips for Success in this Course

1. Get help early on if you are having difficulties. If my office hours don't work for you, we can work something out.
2. Get to know others in the class. Help each other out.
3. If I give bonus opportunities, take advantage of them.
4. If I give study guides, take advantage of them.
5. If a book is required, get the book and use it.
6. Your goal should not be to pass: shoot for an A.
7. If I give a writing assignment with a rubric attached. Use this rubric because this is what I'm looking for.
8. If I give a writing assignment, don't hesitate to get help.
9. Be open-minded. I understand that this class may not be within your subject of interest, but that doesn't mean you can't take interest. It's easier to learn something you have an interest in.

Reading list for online discussion:

Books:

1: Piyushimita Thakuriah, Nebiyou Tilahun, Moira Zellner. Seeing Cities Through Big Data (SCTBD)

1. Aug. 24:

SCTBD: Introduction to Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics (page 1-9)

Kontokosta, C. E. Urban Informatics in the Science and Practice. Journal of Planning Education and Research, 2018, 1-14. https://doi.org/10.1177/0739456X18793716

2. Aug. 31:

SCTBD: Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery (page 11-48)

3. Sep. 7:

SCTBD: The Potential for Big Data to Improve Neighborhood-Level Census Data (page 99-112)

4. Sep. 14:

SCTBD: How Should Urban Planners Be Trained to Handle Big Data? (page 208-217)

5. Sep. 21:

SCTBD: Using an Online Spatial Analytics Workbench for Understanding Housing Affordability (page 233-256)

6. Sep. 28 and Oct. 5:

SCTBD: Modeling Taxi Demand and Supply in New York City Using Large-Scale Taxi GPS Data (Page 405-425)

7. Oct. 19:

Nilsson, P. (2014). Natural amenities in urban space–A geographically weighted regression approach. Landscape and Urban Planning, 121, 45-54.

8. Oct. 26:

Seabrook, N. R. (2009, July). The Obama effect: Patterns of geographic clustering in the 2004 and 2008 presidential elections. In The Forum (Vol. 7, No. 2). De Gruyter.

9. Nov. 2:

SCTBD: Using User-Generated Content to Understand Cities (page 49-64)

10. Nov. 9:

Plunz, R. A., Zhou, Y., Vintimilla, M. I. C., Mckeown, K., Yu, T., Uguccioni, L., & Sutto, M. P. (2019). Twitter sentiment in New York City parks as measure of well-being. Landscape and urban planning, 189, 235-246.

	Module	Week	Topics	Friday exercise
1	Data preps	24-Aug	Introduction to data science	
2		31-Aug	Introduction to R and data type	AS 1: Basic R functions
3		7-Sep	Programming in R loops and functions	AS 2: Build your own functions
4		14-Sep	Preps: Data wrangling	AS 3: Data transformation in R
5	Data visulization and stats	21-Sep	Data visualization and maps	AS 4: Graphics for communication
6		28-Sep	Statistics models	AS5: Solve stats models in R
7		5-Oct		
8		12-Oct	Exam + Project proposal	
9	Application	19-Oct	Geo spatial analysis	AS 6: Vector-Based Analysis in R
10		26-Oct	Spatial autocorrelation and pattern analysis	AS 7: Explore patterns in R
11		2-Nov	Text analytics1	AS 8: Tidy text format
12		9-Nov	Text analytics2	AS 9: Sentimental analysis and connections between words
13		16-Nov	Work on project	
14	Final	30-Nov	Final project presentation	
15		7-Dec	Final project	
16		11-Dec	Final exam week	