UP 494 Adv. Data Science for Planners

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

Online Lectures: Tuesday 3:30 PM - 4:50 PM, click here to join

In-person lab session: Thursday 3:30 PM - 4:50 PM, Oregon lab, 901 W. Oregon St.

Office Hours:

Monday 9:00 am – 11:00am in zoom, click here to join

Thursday 9:00 am – 11:00am in person, TBH 232

TA: Srirang Sohoni (ssohon3@illinois.edu) Office hour: Wed 10:00-12:00

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, machine learning, 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
- Collect, 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 lectures and all types of contents, and 1-2 hours for labs exercise per week.
- All the assignments, exams, discussions, and labs are mandatory.
- In-person lab sessions are delivered every Thursday 3:30 PM - 4:50 PM in Oregon lab
- On campus instruction under COVID: For in person meetings, in order to implement COVID-19-related guidelines and policies affecting university operations, **students will be asked to show their Building Access Status in the Safer Illinois app or the Boarding Pass**. Student can enter the classroom only with a granted building access status. All students, faculty, staff, and visitors are required to wear face coverings in classrooms and university spaces. This is in accordance with CDC guidance and University policy and expected in this class. Following University policy, all students are required to engage in appropriate behavior to protect the health and safety of the community. Students are also required to follow the campus COVID-19 protocols. Students who feel ill must not come to class. In addition, students who test positive for COVID-19 or have had an exposure that requires testing and/or quarantine must not attend class. The University will provide information to the instructor, in a manner that complies with privacy laws, about students in these latter categories. These students are judged to have excused absences for the class period and should contact the instructor via email about making up the work. Students who fail to abide by these rules will first be asked to comply; if they refuse, they will be required to leave the classroom immediately. If a student is asked to leave the classroom, the non-compliant student will be judged to have an unexcused absence and reported to the Office for Student Conflict Resolution for disciplinary action. Accumulation of non-compliance complaints against a student may result in dismissal from the University.

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 Canvas submission. Unless otherwise stated, the lab assignments are due on 3:30 PM of the Thursday that one week after they are assigned (e.g. a lab assigned on Aug. 26st will due on 3:30 PM of Sep. 2nd). You should submit your assignment to Canvas 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 Canvas site. **For any technical issues in Canvas/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.

Final project

All the student needs to finish a final project report with a presentation. The details will be posted on Canvas site 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 8th**.
- The graduate students must finish the project individually.

A project proposal is due by **11:59PM, Oct 19th**. A final report is required as delivery by **11:59 PM, Dec 8th**. The details and requirements will be posted later in Canvas site.

Weekly reading discussion:

This is an initial content engagement discussion forum. You are invited to think about what you already might know about a new idea, concept, problem or closely related concept about urban data science. You need first finish the readings below and submit at least one initial post and two response posts. Some suggested questions (but not limited to) will be available for you to answer.

You should finish reading the required articles **before each lecture starts on Tue 3:30 pm** in order to participate in the in-class discussion. Your timely online posts and reading reflections (**due on each Tursday before lab starts**) in Canvas are required, which worth 100 points total (10*10). Each reading assignment is worth 10 points: 6 for your initial post and 2 for each response post. Any plagiarism is found in any posts will receive a "0".

Initial Post

Your initial post is your opportunity to engage with the prompt in a way that is unique to you. In composing your response, consider how your individual experiences influence your take on the prompt and the course material or articles covered during this module.

An acceptable initial post must meet the following requirements:

- Include at least 8 sentences, excluding any references.
- You are encouraged to 1) study with other students together 2) check out other articles of publications. However, this should never involve 1) one student having possession of a copy of all or part of posts done by someone else; 2) using or copying and pasting others' published and unpublished sentences or words and presenting them as new and original.

Response Posts

Post at least 2 responses in the same thread. Your replies should stimulate more in-depth discussion about the topic. Some ways to accomplish that include:

- Clarify and/or extend your peers' line of thinking.
- Compare/contrast their views on the topic with your own.
- Suggest/question what explanation(s) you think your peers might be missing that could strengthen their arguments.
- End your response with a question to further the dialogue.

Your response posts should meet the following requirements:

- Include at least 50 words, excluding references.
- Use of appropriate evidence from the readings and lessons to support your claims and judgments.
- Any simply or low-quality replies e.g. “I agree with Andrew” or “The article is very interesting” will NOT be accepted.

Grade Point Distribution:

Assignments*10	500 Points Total
Mid-term take home Exam	200 Points
Final project	200 Points (20 proposal, 30 presentation, 150 report)
Participation and in-class activity	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.

Run > Hide > Fight

Emergencies can happen anywhere and at any time. It is important that we take a minute to prepare for a situation in which our safety or even our lives could depend on our ability to react quickly. When we're faced with almost any kind of emergency – like severe weather or if someone is trying to hurt you – we have three options: Run, hide or fight.

Run

Leaving the area quickly is the best option if it is safe to do so.

- Take time now to learn the different ways to leave your building.
- Leave personal items behind.
- Assist those who need help, but consider whether doing so puts yourself at risk.
- Alert authorities of the emergency when it is safe to do so.

Hide

When you can't or don't want to run, take shelter indoors.

- Take time now to learn different ways to seek shelter in your building.
- If severe weather is imminent, go to the nearest indoor storm refuge area.
- If someone is trying to hurt you and you can't evacuate, get to a place where you can't be seen, lock or barricade your area if possible, silence your phone, don't make any noise and don't come out until you receive an Illini-Alert indicating it is safe to do so.

Fight

As a last resort, you may need to fight to increase your chances of survival.

- Think about what kind of common items are in your area which you can use to defend yourself.
- Team up with others to fight if the situation allows.
- Mentally prepare yourself – you may be in a fight for your life.

Please be aware of people with disabilities who may need additional assistance in emergency situations.

Other resources

- **police.illinois.edu/safe** for more information on how to prepare for emergencies, including how to run, hide or fight and building floor plans that can show you safe areas.
- **emergency.illinois.edu** to sign up for Illini-Alert text messages.
- **Follow the University of Illinois Police Department** on Twitter and Facebook to get regular updates about campus safety.

Reading list:

Books:

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

1. **Aug. 30:**
   - 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
   - SCTBD: Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery (page 11-48)
2. **Sep. 6:**
   - SCTBD: The Potential for Big Data to Improve Neighborhood-Level Census Data (page 99-112)
   - Fiack, D., Cumberbatch, J., Sutherland, M., & Zerphey, N. (2021). Sustainable adaptation: Social equity and local climate adaptation planning in US cities. Cities, 115, 103235.
3. **Sep. 13:**
   - SCTBD: How Should Urban Planners Be Trained to Handle Big Data? (page 208-217)
   - Kontokosta, C. E., & Johnson, N. (2017). Urban phenology: Toward a real-time census of the city using Wi-Fi data. Computers, Environment and Urban Systems, 64, 144-153.
4. **Sep. 20:**
   - SCTBD: Using an Online Spatial Analytics Workbench for Understanding Housing Affordability (page 233-256)
   - Bevilacqua, C., Ou, Y., Pizzimenti, P., & Anversa, G. (2020, May). Contextualizing Transition: A Multiscale Approach to Making Resilience-Oriented and Place-Sensitive Strategies. In INTERNATIONAL SYMPOSIUM: New Metropolitan Perspectives (pp. 47-67). Springer, Cham.
5. **Sep. 27:**
   - SCTBD: Modeling Taxi Demand and Supply in New York City Using Large-Scale Taxi GPS Data (Page 405-425)
   - Bischoff, K. (2008). School district fragmentation and racial residential segregation: How do boundaries matter?. Urban Affairs Review, 44(2), 182-217.
6. **Oct. 4:**
   - Kruse, J., Kang, Y., Liu, Y. N., Zhang, F., & Gao, S. (2021). Places for play: Understanding human perception of playability in cities using street view images and deep learning. Computers, Environment and Urban Systems, 90, 101693.
   - Kleine Deters, J., Zalakeviciute, R., Gonzalez, M., & Rybarczyk, Y. (2017). Modeling PM2. 5 urban pollution using machine learning and selected meteorological parameters. Journal of Electrical and Computer Engineering, 2017.
7. **Oct. 18:**
   - Nesbitt, L., Meitner, M. J., Girling, C., Sheppard, S. R., & Lu, Y. (2019). Who has access to urban vegetation? A spatial analysis of distributional green equity in 10 US cities. Landscape and Urban Planning, 181, 51-79.
   - Feng, C., & Jiao, J. (2021). Predicting and mapping neighborhood-scale health outcomes: A machine learning approach. Computers, Environment and Urban Systems, 85, 101562.

8. **Oct. 25:**
    - 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.
    - Maroko, A. R., Maantay, J. A., Sohler, N. L., Grady, K. L., & Arno, P. S. (2009). The complexities of measuring access to parks and physical activity sites in New York City: a quantitative and qualitative approach. International journal of health geographics, 8(1), 1-23.
    - Oakley, D. (2008). Locational patterns of low-income housing tax credit developments: A sociospatial analysis of four metropolitan areas. Urban Affairs Review, 43(5), 599-628.
9. **Nov. 1:**
    - SCTBD: Using User-Generated Content to Understand Cities (page 49-64)
    - Steiger, E., Westerholt, R., Resch, B., & Zipf, A. (2015). Twitter as an indicator for whereabouts of people? Correlating Twitter with UK census data. Computers, environment and urban systems, 54, 255-265.
10. **Nov. 8:**
    - 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.
    - Abdul-Rahman, M., Chan, E. H., Wong, M. S., Irekponor, V. E., & Abdul-Rahman, M. O. (2021). A framework to simplify pre-processing location-based social media big data for sustainable urban planning and management. Cities, 109, 102986.

Fall 2021 Department of Urban and Regional Planning

	Module	Week	Topics	Assignment
1	Data preps	23-Aug	Introduction to data science	R setup
2		30-Aug	Data science basics	AS 1: Basic R functions
3		6-Sep	Loops and functions	AS 2: Build your own functions
4		13-Sep	Data Wrangling	AS 3: Data transformation in R
5	Exploratory data analysis	20-Sep	Data Visualization	AS 4: Graphics for communication
6		27-Sep	Regression Models	AS5: Solve stats models in R
7		4-Oct	Classification: Basic machine learning in urban data analysis	AS6: Apply machine learning algorithms
8		11-Oct	Exam + Project proposal	
9	Application	18-Oct	Spatial data representation	AS 7: Vector-Based Analysis in R
10		25-Oct	Spatial autocorrelation and pattern analysis	AS 8: Explore patterns in R
11		1-Nov	Text data collection and regular expressions	AS 9: Tidy text format
12		8-Nov	Text data on sentiment analysis	AS 10: Sentimental analysis and connections between words
13		15-Nov	Work on project	
14	Final	22-Nov	Fall break	
15		29-Nov	Final project presentation	
16		6-Dec	Final exam week	