UP 317 Introduction to Urban Data Science

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

Lectures: Tuesday 11:00-12:20

labs: Thursday 11:00-12:20 122 1203 1/2 W Nevada

Office Hours:

Wed 10:00 am – 12:00 pm TBH 232, or by appointment

TA: Aakanksha Kulkarni (ask13@illinois.edu) Office hour: Wed 1:00 pm to 3:00 pm TBH 224

Course Overview:
Data science is a multi-disciplinary field that involves scientific methods, processes, algorithms, and systems to extract and further understand knowledge or insights from all kinds of data. Especially for planners or social scientists, they use a wide range of data from different sectors such as transportation, housing, physical environment to understand the complex urban issues. This is the entry-level data science course for undergraduate students in urban planning. You will learn a set of fundamental concepts, skills, and tools in R for effective data analysis. We will start with basic data import, data cleansing/transformation, and will introduce data visualization later for communication purposes especially for planners. This course builds a common foundation for quantitative analysis among undergraduate and graduate students for a wide application in one or more domain-specific courses in their capstone/thesis/dissertation work in the future. No previous coding experiences are required.

This syllabus is subject to change by the instructor.

Course Outcomes:

1. Understand the basic concepts and workflows in data science
2. Collect, import, tidy, export, and manipulate data effectively and efficiently.
3. Analyze and interpret data using R in the urban planning discipline
4. Apply the fundamental and basic quantitative techniques in social research.
5. Produce and interpret professional plots, graphics, and maps using R in the urban planning discipline.
6. Interpret and summary results professionally for communication

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 to 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. 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, assignments, and labs.

- Each student is expected to devote 1-2 hours per week learning the lecture contents, and 1-2 hours for lab exercise per week.
- **Engagement**: The engagement is demonstrated in various ways for lectures and labs: e.g. in-class discussion, in-class exercise, in-class group work, in-class presentation, reading discussion, short essay summary etc. The in-class activities for online lectures will be posted via these platforms but not limited to: 1) shared google documents/sheets 2) Google Jamboard; 3) Padlet online. Bonus credits will be allotted for some of the in-class exercise: either essay questions for reading discussions, or mini project, etc.
- Students are also expected to complete the assigned readings prior to class and to come to lectures prepared for thoughtful participation and discussion.
- All the assignments, exams, and labs are mandatory.

Required Textbook

- R for Data Science import, tidy, transform, visualize and model data. By Hadley Wickham & Garrett Grolemund

Software

Students can install RStudio and R on their personal computers for free through https://rstudio.com/products/rstudio/download/

Lab Assignments and Late Work Policy:

In addition to the exams, you will be asked to complete 11 lab assignments. Assignments must be turned in via Canvas site submission. **You will receive a zero on the assignment if it is not submitted. Note the lowest grade among the 11 assignments will be dropped.** Unless otherwise stated, the lab assignments are due on 11 AM of the Thursday that one week after they are assigned (e.g. a lab assigned on Sep 3rd will due on 11 AM of Sep. 10th). You should submit your assignment to Canvas website. An assignment**, including lab assignments, mid-term exam, 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. **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.

Exams:

The exam will take place in Canvas with one-week period (~90min). The exam will be a combination of multiple-choice, true-and-false, and short answer questions. Instead of regurgitating facts, my tests are designed so that you think about the key concepts of the topics we have covered. The exam cannot be re-taken. **Consistent with UIUC guidelines, if you cannot take a regularly scheduled exam because of authorized University activities, you will have the opportunity to take a make-up exam at an alternate time. Make-up exams for absences due to any other reason will be at the discretion of the instructor. You must notify me beforehand if you need 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. The details will be posted on Canvas later this semester.

The final project can be finished as a group. You should email the instructor about your group info by **11:59 PM, Oct 1st**.

A project proposal is due by **11:59PM, Oct 13th**. A final project presentation is needed in the week of Dec 4th.

Grade Point Distribution:

Assignments*10	50 Points Each, 500 Points Total
Mid-term Exam	200 Points
Final project	200 Points (proposal 50 points, presentation 150 points)
Participation and in-class activity	100 Points
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 more quickly, 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 of when a discussion or assignment closes. If you would like to get help on an assignment ahead of the deadline, please email me! I'm happy to give preliminary feedback or answer 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 Succeeding in this Course

1. Get help early on if you are having difficulties. Come to my office if you need to. 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 us it.
6. Your goal should not be to pass; shoot for an A.
7. If I give a writing assignment it will have 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:

1. **Aug. 28:**
   - 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
   - Wang, S. J., & Moriarty, P. (2018). The Urgent Need for Advancing Urban Sustainability. In Big Data for Urban Sustainability (pp. 1-21). Springer, Cham.
2. **Sep. 4:**
   - Vermiglio, C., Kudo, H., & Zarone, V. (2020, May). Making a Step Forward Towards Urban Resilience. The Contribution of Digital Innovation. In INTERNATIONAL SYMPOSIUM: New Metropolitan Perspectives (pp. 113-123). Springer, Cham.
3. **Sep. 11:**
   - Wang, S. J., & Moriarty, P. (2018). The potential for big data for urban sustainability. In Big data for urban sustainability (pp. 45-63). Springer, Cham.
4. **Sep. 18:**
   - Wang, S. J., & Moriarty, P. (2018). Barriers to the Implementation of Big Data. In Big Data for Urban Sustainability (pp. 65-80). Springer, Cham.
5. **Sep. 25:**
   - Buffoli, M., Rebecchi, A., Dell'Ovo, M., Oppio, A., & Capolongo, S. (2020). Transforming the Built Environment Through Healthy-Design Strategies. New Metropolitan Perspectives, 187.
6. **Oct. 2:**
   - Monteiro, C. S., Costa, C., Pina, A., Santos, M. Y., & Ferrão, P. (2018). An urban building database (UBD) supporting a smart city information system. Energy and Buildings, 158, 244-260.
7. **Oct. 16:**
   - Portnov, B. A. (2020). On Locational Commonalities of Largest Cities Worldwide. Innovations in Urban and Regional Systems; Springer: Cham, Switzerland, 23-48.
8. **Oct. 23:**
   - Graziano, M., Alexander, K. A., Liesch, M., Lema, E., & Torres, J. A. (2019). Understanding an emerging economic discourse through regional analysis: Blue economy clusters in the US Great Lakes basin. Applied Geography, 105, 111-123.
   - Bounoua, L., Nigro, J., Zhang, P., Thome, K., & Lachir, A. (2018). Mapping urbanization in the United States from 2001 to 2011. Applied geography, 90, 123-133.
9. **Oct. 30:**
   - Desjardins, M. R., Hohl, A., & Delmelle, E. M. (2020). Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. Applied Geography, 118, 102202.
10. **Nov. 6:**
    - Shelton, T. (2021). Gameday homes: Mapping emerging geographies of housing speculation and absentee ownership in the American South. Cities, 115, 103230.
11. **Nov. 13:**
    - Luo, F., Cao, G., Mulligan, K., & Li, X. (2016). Explore spatiotemporal and demographic characteristics of human mobility via Twitter: A case study of Chicago. Applied Geography, 70, 11-25.

	Week	Topics	
1	21-Aug	Intro to Data Science	R setup
2	28-Aug	Introduction to R	Assignment 1 Basics and data type
3	4-Sep	Basic control structures	Assignment 2 Loops and functions
4	11-Sep	Data import	Assignment 3 Get your data ready
5	18-Sep	Tidy messy data	Assignment 4 Evaluate air quality in Sydney
6	25-Sep	Data Manipulation	Assignment 5 Analysis population based on American Community Survey
7	2-Oct	Relational data	Assignment 6 Education level and housing/income analysis
8	9-Oct	Review & Midterm exam	
9	16-Oct	Data summaries	Assignment 7 Summary Statistics
10	23-Oct	Basic Graphics design	Assignment 8 Visualize demographic information in midwest counties
11	30-Oct	Data virsualization	Assignment 9 Visualize air quality in Sydney
12	6-Nov	Data visualization for spatial data	Assignment 10 Draw maps using R
13	13-Nov	Working with time series data	Assignment 11 Understand flights in NYC,2013
14	20-Nov	Fall break	
15	27-Nov	Group work	
16	4-Dec	Final exam week	