Course Description

This course introduces the techniques and tools used to assemble, manage, and analyze many types of data used to support decision-making in urban environments. We will work with real-world datasets. Since these datasets are often messy and incomplete, an important part of the course will be learning to clean and analyze imperfect data. The goal of this course is to prepare students to apply urban data analytics in practice. More emphasis will be placed on practical applications of urban data analytics than on the theory underlying these methods. Students will also learn good coding practices for managing urban data science projects, including version control, documentation, and modular design.

The class has no formal prerequisites, and no programming experience is required. Undergraduate students at all levels are welcome to take this course.

Students should have laptop computers that they bring to class. Ideally, their computers will have sufficient memory and processing power to run R, RStudio, and QGIS with medium-sized datasets (any recent Mac or PC should be sufficient). Students should install these (free) programs before the third class. Students working on Windows machines should install git as well, either directly from Git, or by installing the Windows Subsystem for Linux (if you’re not sure which you should install, you should probably install directly from Git).

If you do not have a laptop you can bring to class, or if you have a Chromebook or other system unable to run these applications, please contact me as soon as possible to discuss options.

Learning Objectives

By the end of this course, students will be able to

[1] Think about data and problem-solving using algorithms
[2] Write understandable and reproducible code to analyze urban data
[3] Manage that code in a version control system
[4] Perform basic exploratory urban data analysis
[5] Visualize urban data, including spatial data
[7] Scrape structured and unstructured information from the internet
[8] Perform text analysis to make sense of unstructured text data
[9] Work with GIS datasets in a statistical programming environment

Major and Minor Requirements
This course fulfills the following curriculum requirements
- Urban Studies and Planning minor: this course can count as one of the 3 required PLAN electives
- Data Science minor: this course can be used as an elective

Course Materials
All required materials will be available on Sakai. There are no textbooks for this course. Readings should be completed before the class they are listed with.

Class Format
Class meetings will consist of lectures, hands-on exercises, and discussion. Please come to class ready to discuss the week’s homework and any issues encountered. Hands-on exercises are critical to promote understanding of data science techniques and to troubleshoot inevitable issues. Students should bring their laptops to class; if this is hardship please contact the instructor immediately.

Course schedule
- **January 10**: Introduction to R/QGIS/git, assist with installation and setup on student machines as needed.
- **January 12**: Data science ethics
  - **Readings** (all on Sakai under Readings -> Ethics):
  - Green, *The Smart Enough City*, ch. 1. (MIT Press, 2019)
  - *No class January 17 due to MLK Jr. day holiday*
- **January 19–24**: Git and GitHub
  - Homework #1 due Jan 27
- **January 26–February 9**: Exploratory data analysis and visualization; applications of good coding practices and version control. Data cleaning.
  - Homework #2 due Feb 15
- **February 14–23**: Spatial data. Introduction to using spatial data in R and QGIS.
  - Homework #3 due Mar 1
- **February 28–March 9**: Web scraping. Retrieving structured and unstructured data from web sites.
  - *Submit choice of analysis for critique by Mar 1*
  - Homework #4 due Mar 22
- **March 21–24**: Text analysis. Use R to perform sentiment analysis or other text analysis of data scraped from websites or retrieved using APIs.
  - Homework #5 due Mar 29
  - *Urban Data Analysis Critique due Mar 29*
- **March 28–April 6**: Networks. Analyze and visualize data on origins, destinations, and flows using R and QGIS.
  - Homework #6 due Apr 12
Final Project Proposal due Apr 5

- **April 11–13:** Raster datasets. Use raster spatial data in R and QGIS.
  Homework #7 due Apr 19
- **April 18–25:** Open lab to work on final projects.
- **April 27 & 29:** Final project presentations. *Note: Class meets Fri, April 29, 4-7 PM*

Final Project Report due Apr 28

Assignments and grading

Your course grade will consist of:

- **7 homework exercises**
  - Homework exercises are generally due every 2 weeks on Tuesday, and are based on the weeks’ lecture topic. All homework assignments will be posted at least 2 weeks before the due date.

- **One critique of an existing urban data analysis**
  - You will find an existing urban data analysis, and critically evaluate it. What was the research question? Were the data and methods appropriate to address this question? How do you think this analysis should impact policy (if at all)?
  - You may find your analysis in the popular press, government reports, or the academic literature, but it should contain enough detail to evaluate the methods and data used.
  - You should submit your choice of analysis to me before you start, to ensure it fits the criteria for the assignment.

- **One independent project**
  - The final assignment is an independent project analyzing data you find to answer an urban-related question you develop.
  - This project will consist of a proposal, a final report (8-15 pages double spaced), and a 10-minute in-class presentation.
  - The project should use an existing dataset, analyzed using methods discussed in this class.
  - The project should include a discussion of policy implications of your results.
  - Before you submit your proposal, you must meet with me (in person or on Zoom) to discuss the project. You may either visit my office hours or set up an appointment at another time. You are of course welcome to meet with me as often as needed as you develop and complete the project.

Each graded component will contribute to the final grade as follows:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Due (all at 11:59 PM)</th>
<th>% Grade</th>
<th>Final homework combined = 50%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Homework 1: Version control</td>
<td>Jan 27 (Thursday)</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>2 Homework 2: Exploratory data analysis</td>
<td>Feb 15</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>3 Homework 3: Spatial data</td>
<td>Mar 1</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>4 Homework 4: Web scraping</td>
<td>Mar 22</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>5 Homework 5: Text analysis</td>
<td>Mar 29</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>6 Homework 6: Network analysis</td>
<td>April 12</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>7 Homework 7: Raster data analysis</td>
<td>Apr 26</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>8 Urban Data Analysis Critique</td>
<td>Mar 29</td>
<td>20%</td>
<td></td>
</tr>
<tr>
<td>9 Final Project: Proposal</td>
<td>Apr 5</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Final Project: Report</td>
<td>Apr 28</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Final Project: Presentation</td>
<td>Apr 27/29</td>
<td>10%</td>
<td></td>
</tr>
</tbody>
</table>

Final grades will be calculated as follows:
Late assignments

To keep the course on schedule, it is important that assignments be submitted on time. I understand that circumstances may occasionally arise that prevent you from turning in assignments on time. Thus, you may submit one assignment (other than the final project) up to one week late without penalty, although I would appreciate an email before the due date letting me know you will be turning it in late. Additional late assignments will receive a 10% deduction for each day late up to 5 days, after which they will not be accepted. (If you turn in multiple assignments late, I will apply the one-week grace period to whichever assignment will result in you having the highest grade.) Exceptions to this late work policy may be made on a case-by-case basis—please contact me as soon as possible if you anticipate needing an exception.

If you encounter technical difficulties that may prevent you from completing an assignment on time, please contact me as soon as possible for assistance. Depending on the nature of the problem, deadline extensions may be approved on a case-by-case basis. Approved deadline extensions due to technical difficulties will incur no late penalties and will not count as a late assignment.

University resources

In addition to class time and office hours, there are several resources on campus that may help you with the concepts and assignments in this course, and I encourage you to use them.

Davis Library Research Hub
https://library.unc.edu/data/

The Research Hub is a service provided by UNC Libraries to assist students with questions about R, GIS, and statistics generally. The service is currently operating via video calls and online chat.

Odum Institute Online Statistics Help Desk
https://odum.unc.edu/statistics-help-desk/

The Odum Institute provides online statistical support for R and statistics, generally for more advanced questions than the Research Hub can address. This may be useful when working on your final project.

Writing Center
writingcenter.unc.edu

The Writing Center provides one-on-one assistance with writing, editing, and proofreading; this may also be helpful for your final project.

Absences

I understand there are times when you may need to miss class for any number of reasons. If you are going to miss class, I would appreciate an email letting me know. I will attempt to work with you so you can keep up if you must miss class for due to reasonable circumstances (e.g. illness, emergency, job interview).

The University has a policy to request University Approved Absences (described below), but it is not necessary to do so—an informal email to me is fine. If you are ill or think you may have COVID-19, please do not come to class—let me know and we can make alternate arrangements.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Grade</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Grade</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
<th>Grade</th>
<th>Lower Boundary</th>
<th>Upper Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
University Approved Absences are defined by the university at attendance.unc.edu:

- Authorized University activities
- Disability/religious observance/pregnancy, as required by law and approved by Accessibility Resources and Service and/or the Equal Opportunity and Compliance Office (EOC)
- Significant health condition and/or personal/family emergency as approved by the Office of the Dean of Students, Gender Violence Service Coordinators, and/or the Equal Opportunity and Compliance Office (EOC).

You can request a University Approved Absence through the University Approved Absence Office by using this request form.

Mask use
This semester, while we are in the midst of a global pandemic, all enrolled students are required to wear a mask covering your mouth and nose at all times in our classroom. This requirement is to protect our educational community—your classmates and me—as we learn together. If you choose not to wear a mask, or wear it improperly, I will ask you to leave immediately, and I will submit a report to the Office of Student Conduct. Students who have an authorized accommodation from Accessibility Resources and Service have an exception. For additional information, see Carolina Together.

Spring 2022 course delivery
As long as it is possible to do so safely, we will be meeting in person this semester. I understand the ongoing COVID-19 pandemic may require changes to this plan and will be monitoring the situation closely. If I need to change the format of the course temporarily due to outbreaks of illness, I will announce this via email and the course Sakai site.

Honor code
All students are expected to follow the guidelines of the UNC honor code. In particular, students are expected to refrain from “lying, cheating, or stealing” in the academic context. If you are unsure about which actions violate the honor code, please see me or consult honor.unc.edu.

I will report any honor code violation to the Office of Student Conduct.

For this course:
- You are permitted and encouraged to seek advice and suggestions from other class members on the written assignments and homework, unless specifically instructed otherwise. This may include exchanging drafts for feedback and/or proofreading. Everything you turn in and all analysis must be your own work.
- In all written and presented work, you must cite or otherwise fully attribute all ideas, data, and other information that are not your own. This includes information presented in tables, graphs, appendices, etc.

Title IX resources
The University is committed to providing an inclusive and welcoming environment for all members of our community and to ensuring that educational and employment decisions are based on individuals’ abilities and qualifications. Consistent with this principle and applicable laws, the University’s Policy Statement on Non-Discrimination offers access to its educational programs and activities as well as employment terms and conditions without respect to race, color, gender, national origin, age, religion, creed, genetic information, disability, veteran’s status, sexual orientation, gender identity or gender expression. Such a policy ensures that only relevant factors are considered and that equitable and consistent standards of conduct and performance are applied.
If you are experiencing harassment or discrimination, you can seek assistance and file a report through the Report and Response Coordinators (see contact info at safe.unc.edu) or the Equal Opportunity and Compliance Office, or online to the EOC at https://eoc.unc.edu/report-an-incident/.

Diversity statement
I value the perspectives of individuals from all backgrounds reflecting the diversity of our students and my goal is to create a safe space for everyone in this class. I broadly define diversity to include race, gender identity, national origin, ethnicity, religion, social class, age, sexual orientation, political background, and physical and learning ability. Please let me know if there is anything I can do to improve, I appreciate suggestions.

Accessibility resources and services
The University of North Carolina at Chapel Hill facilitates the implementation of reasonable accommodations, including resources and services, for students with disabilities, including mental health disorders, chronic medical conditions, a temporary disability or pregnancy complications resulting in barriers to fully accessing University courses, programs and activities.

Accommodations are determined through the Office of Accessibility Resources and Service (ARS) for individuals with documented qualifying disabilities in accordance with applicable state and federal laws. See the ARS Website for contact information: https://ars.unc.edu or email ars@unc.edu.

Counseling and Psychological Services
CAPS is strongly committed to addressing the mental health needs of a diverse student body through timely access to consultation and connection to clinically appropriate services, whether for short or long-term needs. Go to their website: https://caps.unc.edu/ or visit their facilities on the third floor of the Campus Health Services building for a walk-in evaluation to learn more. Students can also call CAPS 24/7 at 919-966-3658 for immediate support.

Resources
Our purpose as professors is to help you to excel in this learning environment. Should you need further assistance beyond the help of the professor, please consult the following on-campus resources:

- **The Writing Center** provides one-on-one assistance to students. To make an appointment, browse the Writing Center’s online resources, or submit a draft online. They have additional useful information, such as handouts on how to cite online. writingcenter.unc.edu
- **The Learning Center** offers individual consultations, peer tutoring, academic coaching, test prep programming, study skills workshops, and peer study groups. learningcenter.unc.edu
- **Campus Health** provides ambulatory primary medical care, mental health services and wellness programs along with selected specialty services. campushealth.unc.edu

Acceptable use policy
By attending the University of North Carolina at Chapel Hill, you agree to abide by the University of North Carolina at Chapel Hill policies related to the acceptable use of IT systems and services. The Acceptable Use Policy (AUP) sets the expectation that you will use the University’s technology resources responsibly, consistent with the University’s mission. In the context of a class, it’s quite likely you will participate in online activities that could include personal information about you or your peers, and the AUP addresses your obligations to protect the privacy of class participants. In addition,
the AUP addresses matters of others’ intellectual property, including copyright. These are only a couple of typical examples, so you should consult the full Information Technology Acceptable Use Policy, which covers topics related to using digital resources, such as privacy, confidentiality, and intellectual property.

Additionally, consult the University website “Safe Computing at UNC” for information about the data security policies, updates, and tips on keeping your identity, information, and devices safe.

**Syllabus changes**
The professor reserves the right to make changes to the syllabus, including project due dates and test dates. These changes will be announced as early as possible.

**Grade appeal process**
If you feel you have been awarded an incorrect grade, please discuss with me. If we cannot resolve the issue, you may talk to our departmental director of undergraduate studies or appeal the grade through a formal university process based on arithmetic/clerical error, arbitrariness, discrimination, harassment, or personal malice. To learn more, go to the Academic Advising Program website.