CSC 110 - Data Science and Society - Spring 2022

Course Information

Course Location	Meeting Days	Time
Watson 132	Tuesdays & Thursdays	9:50-11:05am

Instructor Information

Instructor Information	Office Location	Hours
Dr. Jason S. Byers jabyers@davidson.edu (704) 894-2760	Main Office Chambers 2256 Data CATS, drop-in hours Hunt Hub	TR: 8-10am TBD

Syllabus

Course Home
Everything you need for this class (announcements, resources, assignments and other activities) will be posted on the course website. Please plan to check the page regularly.

Course Meeting Link
TBD

Course Description
Data plays a central and important role in modern society. Nearly every web search, phone call, transaction, and countless other activities are tracked, analyzed, and used to predict future actions. Data science is a collection of analytical and computational methods to enable insight, understanding and predictions to be drawn from data. This course provides an introduction to the methods of data science, including visualization, manipulation, programming and modeling, in the context of social justice. All work will be done in R, a freely available data analysis software.

Learning Outcomes
Together, we will strive for your individual and collective success in achieving the learning outcomes of this course. At the conclusion of this course, students will be able to:

- Define and describe the varied nature of data and data science
- Describe and discuss ethical issues in the practice of data science
- Organize and clean data using R
- Visualize data effectively and persuasively with data, using accurate, unbiased and aesthetically pleasing visualizations
- Apply data science methods to illuminate and analyze issues of injustice or structural inequality

Curricular Connections
This course satisfies the graduation requirement in Justice, Equality, and Community. It is also an introductory course for the Data Science minor, and satisfies the Mathematical & Quantitative Thought distribution requirement.

Prerequisites
There are no prerequisites for this course. In particular, I will assume no previous experience with R, computer programming or statistics.

Course Materials
To maximize access to this class, we will use freely available textbooks, videos, and other resources, with a focus on the following:

- Primary Text (LOTS): Leading an Opportunity Task Force Report
- Primary Text (R4DS): Wickham, Hadley and Garrett Grolemund. 2016. "R for Data Science: Import Tidy, Transform, Visualize, and Model Data." O'Reilly Media. This book is freely available online. It is also available in paperback, if you prefer a hard copy. Warning: some content and the numbering system differs between print and online versions; I will exclusively refer to the free online version.
- Primary Text (ADSP): Grolemund, Garrett. 2014. "Hands-On Programming with R: Write Your Own Functions and Simulations." O'Reilly Media. This book is freely available online. It is also available in paperback, if you prefer a hard copy. Warning: some content and the numbering system differs between print and online versions; I will exclusively refer to the free online version.
- Primary Text (GCB): Lovelace, Robin, Jakub Nowosad, and James Muenchow. 2019. "Geocomputation with R." CRC Press. This book is freely available online. It is also available in paperback, if you prefer a hard copy. Warning: some content and the numbering system differs between print and online versions; I will exclusively refer to the free online version.
- Primary Text (USDR): Engel, Claudia A. 2019. "Using Spatial Data with R" This book is freely available online.
- Primary Text (MSR): Wickham, Hadley. 2021. "Mastering Shiny: Build Interactive Apps, Reports and Dashboards Powered by R." O'Reilly Media. This book is freely available online.

Reference Texts
- Supplementary Text (DSB): Cetinkaya-Rundel, Mine. 2021. "Data Science in a Box." This book is freely available online.

Software
You will use two freely available programs, R and RStudio, in order to complete the assignments for this course. R and RStudio are installed on all Davidson campus computers. They are also freely available to install on your own computer.

Access and Accommodation
The college welcomes requests for accommodations related to disability and will grant those that are determined to be reasonable and maintain the integrity of a program or curriculum. To make such a request or to begin a conversation about a possible request, please contact the Office of Academic Access and Disability Resources, which is located in the Center for Teaching and Learning in the E.H. Little Library: Beth Bilal, Director, bebilal@davidson.edu, 704-894-2129, or Alysen Beaty, Assistant Director, albeaty@davidson.edu, 704-894-2939. It is best to submit accommodation requests within the drop/add period; however, requests can be made at any time in the semester. Please keep in mind that accommodations are not retroactive.

Course Organization
Modes of learning in this class (whether assessed directly or indirectly) require a range of skills and abilities. Every student's success is important to me, and I am happy to work with you to develop strategies for success in this class. For Fall 2021, we will be meeting both in person and remotely; a key objective to participate fully in the collaborative environment that is necessary to maximize your learning.

Learning R and using it to illuminate and analyze issues of injustice or structural inequality takes regular and repeated practice. This is not a lecture class, but an active learning environment, relying on support from classmates, the embedded tutor and the instructor, and interspersed with discussions and short demonstrations and presentations.

- **In-Class Activities** Each class day will involve a significant amount of hands-on work with R and data. In order to learn from these activities, you must do the assigned readings and videos before you come to class, and be prepared to ask (and answer) questions before diving into a small group discussion or activity.
- **Labs.** Bi-Weekly assignments (due approximately every other Thursday at 5 pm Eastern, for a total of 5 labs) will provide you with regular practice using data science methods in R, and applying these methods to illustrate or advocate for a social justice policy. These assignments will build on the material presented in class, and require you to apply the basic concepts in new ways. All lab work is to be done with a partner live on Zoom. While you may be excited to try something on your own, you should never spend more than 15 minutes working on an assignment by yourself. All code must be the work of you and your partner. You may get help on labs from Data CATS tutors, from the embedded tutor, or from me. You may also search for existing advice on the internet. You may not ask any other person, whether at Davidson or elsewhere (including the Internet) to help you solve a problem.
- **Quizzes.** In general, weeks without labs will have quizzes released on Thursday and due on the following Thursday at 5 pm Eastern, for a total of 5 quizzes. Quizzes cover all course material that has been presented between each quiz, with an emphasis on techniques used in the previous lab. Quizzes are to be completed individually, with no help from anyone, in a limited amount of time. Quizzes are open book, and open notes.
- **Final Group Project.** The goal of the final project is for you to apply the data science skills learned in this course to real data to answer a question that helps illuminate and analyze issues of injustice or structural inequality. You will work in teams of three to build a dashboard with graphics and text explanations and interpretations. Intermediate deadlines will entail submitting your research questions, data sources, design sketch, and initial maps and graphs.

Attendance Policy
Missing class will adversely affect your grade in many ways. In addition, the college attendance policy will be enforced: missing more than 25% of class meetings makes you eligible for a failing grade. Please look carefully at the syllabus during the first week of class. Should there be a conflict between any class session, lecture or assignment due date and a religious holiday or observance, athletic contest, or another academic or personal commitment please let me know well in advance. Religious observance warrants a legitimately excused absence. If you must miss class for any reason, excused or otherwise, you are responsible for getting notes from a classmate and turning in all work on time. Each student will be granted 2 unexcused absences.

Getting Help
It is normal and expected that all students will need help outside of class with the material in this course. Because a language like R is rarely learned with practice, an important source of help is additional exercises, in the required textbook or optional online resources provided on the course web page. The following additional resources are also available.

- **AT Sessions** The AT for this class will attend class and help me answer your questions as they arise. There will also be at least two different AT sessions during the week. All students are strongly encouraged to attend an AT session every week.
- **Office Hours** I welcome you to visit me during the hours listed at the beginning of this document. It is a good practice to make an appointment with me even if outside of the listed hours of availability.
- **Data Cats.** Data CATS offers free assistance to students working on data-focused assignments and projects. Student consultants are highly qualified to help you debug R code, find datasets, perform analyses, and make visualizations. Data CATS is located in the Hunt Hub.
- **Math and Science Center.** The Math & Science Center offers free assistance to students in all areas of math and science, with a focus on the introductory courses. Trained and highly qualified peers hold one-on-one and small-group tutoring sessions on a drop-in basis or by appointment, as well as timely recap sessions ahead of scheduled reviews. Emphasis is placed on thinking critically, understanding concepts, making connections, and communicating effectively, not just getting correct answers. In addition, students can start or join a study group and use the MSC as a group or individual study space. It is located in the Center for Teaching & Learning (CTL) on the first floor of the College Library.
- **Reusing/Sharing Code.** Many of the datasets we will discuss and analyze are publicly available, so they may have been extensively discussed and analyzed. Unless explicitly instructed otherwise, you may use available code and resources for course activities (e.g., GitHub repos, StackOverflow answers) but you must cite the source of the code/resource within your program files and/or document. Recycled code that is discovered that is not properly cited may be considered as plagiarism. When working in groups on class assignments you are welcome to discuss problems together and ask for general advice, but you may not share or use code from another group.
- **Honor Code.** Please adhere to the Davidson College Honor Pledge.

Grading

Category	Points
Attendance	10 Points
Participation	10 Points
Labs	30 Points
Quizzes	30 Points
Final Project	20 Points

Schedule

A tentative class schedule of topics, readings and due dates is available below. Minor adjustments will be made as needed, on the course's web page. Please double check the web page before doing each reading assignment.

Week 1

Topics
- Introduction
- Downloading R/RStudio

Date	Readings
1/25	What is Data Science? How do we Learn?
1/27	HOPR Appendix A Class Notes

Assignments
- Download and install R and RStudio on your personal machines

Week 2

Topics
- Introduction
- Introduction to R and RStudio
- R Markdown
- Getting and Loading Data
- Dealing with Messy Data

Date	Readings
2/1	HOPR Chapters 1-3 R4DS Chapter 4 R4DS Chapter 6 R4DS Chapter 27.1-27.3
2/3	HOPR Chapter 7 R4DS Chapter 12.5

Week 3

Topics
- Data Visualization
- Understanding Economic Mobility

Date	Readings
2/8	R4DS Chapter 3.1-3.4 Data Visualization in R
2/10	R4DS Chapter 3.5-3.10 R4DS Chapter 28

Assignments
- Lab 1 Assigned

Week 4

Topics
- Transforming Data

Date	Readings
2/15	R4DS Chapter 5.1-5.4 LOTF Interactive Summary LOTF Chapter 1
2/17	R4DS Chapter 5.5-5.7

Assignments
- Lab 1 DUE
- Quiz 1 Assigned

Week 5

Topics
- Exploring Data
- Segregation and Education

Date	Readings
2/22	R4DS Chapter 7.1-7.4
2/24	R4DS Chapter 7.5-7.7 R4DS Chapter 12.3-12.7 LOTF Chapter 3

Assignments
- Lab 2 Assigned
- Quiz 1 DUE

Week 6

Topics
- Tidy Data
- Child & Family Stability
- Relational Data

Date	Readings
3/1	R4DS Chapter 9 R4DS Chapter 10.1-10.5 R4DS Chapter 12.1-12.3
3/3	R4DS Chapter 12.4-12.7 R4DS Chapter 13.1-13.4 LOTF Chapter 5

Assignments
- Lab 2 Assigned
- Quiz 2 DUE

Week 7

Topics
- Relational Data
- Group Workshop I

Date	Readings
3/8	R4DS Chapter 13.5-13.7 R4DS Chapter 14.1-14.2 R4DS Chapter 15.1-15.2 R4DS Chapter 16.1-16.2
3/10	**Workshop** Data Sets and Research Questions

Assignments
- Lab 3 DUE
- Quiz 3 Assigned

Week 8

Topics
- Spring Break

Date	Readings
3/15	**No Class**
3/17	**No Class**

Week 9

Topics
- Mapping Geospatial Data
- Segregation

Date	Readings
3/22	USDR Chapter 1 GCR Chapter 1 LOTF Chapter 2
3/24	USDR Chapter 3.1-3.4 GCR Chapter 6.1-8.2

Assignments
- Lab 4 Assigned
- Quiz 3 DUE

Week 10

Topics
- Mapping Geospatial Data
- Shiny II Dashboard

Date	Readings
3/29	USDR Chapter 3.5-3.6 GCR Chapter 8.3-8.6
3/31	MSR Chapter 1-4 MSR Chapter 12

Assignments
- Lab 4 DUE
- Quiz 4 Assigned

Week 11

Topics
- Group Workshop II
- Shiny II Dashboard

Date	Readings
4/5	Group Workshop II
4/7	MSR Chapter 5 MSR Chapter 6

Assignments
- Lab 5 Assigned
- Quiz 4 DUE

Week 12

Topics
- Shiny II Dashboard
- Data Science and Modeling

Date	Readings
4/12	MSR Chapter 7 MSR Chapter 8
4/14	MSR Chapter 9 MSR Chapter 10

Assignments
- Lab 5 DUE
- Quiz 5 Assigned

Week 13

Topics
- Group Workshop III
- Topics in Data Science

Date	Readings
4/19	Group Workshop III
4/21	R4DS Chapter 22 R4DS Chapter 23

Assignments
- Quiz 5 DUE

Week 14

Topics
- Group Workshop IV
- Topics in Data Science

Date	Readings
4/26	R4DS Chapter 24 R4DS Chapter 25
4/28	Group Workshop IV

Week 15

Topics
- Group Presentations

Date	Readings
5/3	Group Presentations
5/5	Group Presentations