

DSC 482/MTH 445

Applied math for data science

Syllabus – Fall 2022

Marcus Birkenkrahe

1. General Course Information

- Meeting Times: Tuesday + Thursday from 8 AM - 9.15 AM
- Meeting place: Lyon Building Computer Lab room 104
- Professor: Marcus Birkenkrahe
- Professor's Office: Derby Science Building 210
- Phone: (870) 307-7254 (office) / (501) 422-4725 (private)
- Office hours: Monday through Friday after class (see Canvas)
- Textbook: The Book of R (Part III) by T Davies ([NoStarch, 2016](#))

2. Objectives

Data science is concerned with getting data to work for us, to give us its (presumed) hidden treasures. Data science has been called "the sexiest job of the 21st century". Even if you don't want to become a professional data scientist, it's helpful to master the basic concepts if you want to succeed in today's data-driven business.

Data science is known to be math and stats heavy, but most of the math power is under the hood - you don't see much of it. This course will help you bridge the gap between what you already know and what is required to do exploratory data analysis. In this class, real datasets are used extensively, and all work is supported by R coding.

We cover applications like PCA, mixture distributions, random graph models, Hidden Markov models, linear and logistic regression, Monte Carlo simulations, and neural networks. The course leads the student to think critically about the how and why of probability and statistics, and to see the big picture. The course is not theorem/proof oriented.

You will get lectures, practice sessions with interactive notebooks (in Emacs) and weekly quizzes. You will work on a term project alone or in a group that you will present. Prerequisites are calculus, some matrix algebra, and some experience in programming (it doesn't need to be R). This course is cross-referenced as a math special topics course (MTH 445). This course will be offered again in fall 2024.

3. Target audience

The course is for anyone who is interested in becoming more data literate in their own field of interest - be it languages, theatre, biology, psychology or exercise science - and growing their personal skill stack.

4. Student Learning Outcomes

Students who complete "Applied math for data science" (DSC 482) will have a foundational understanding and coding abilities in these areas:

- Elementary summary statistics and implementation in R
- Visually exploring data with plotting across a range of plot types
- Overview of the concepts of probability and random variables
- R implementation and stats interpretation of distributions
- Literate programming with Emacs and Org-mode
- Statistical programming with R + ESS
- Professional presentation of agile team projects

In connection with a course on linear algebra, this course prepares participants especially for taking DSC 305, "Machine learning" (spring 2023).

5. Course requirements

Introduction to programming (CSC 100 or CSC 115), calculus, matrix algebra. Some knowledge of, and experience with programming and using the R language is useful but not critical. Curiosity is essential. You will gain data literacy skills by taking this course. The course will prepare you for further studies in computer and data science, or in other disciplines that use modern computing, i.e. every discipline, from accounting to zoology).

6. Grading system

The table lists course requirements, number of units (classroom sessions), points-per-unit (PPU), total points per requirement, and the impact of a requirement on your total grade.

Table 1:

REQUIREMENT	UNITS	PPU	TOTAL	% of TOTAL
Final exam	1	100	100	20.
Home Assignments	15	10	150	30.
Class Assignments	15	10	150	30.
Sprint review	5	20	100	20.
TOTAL			500	100.

You should be able to see your current grade at any time using the gradebook for the course.

7. Grading table

This table is used to convert completion rates into letter grades. For the midterm results, letter grades still carry signs, while for the term results, only straight letters are given (by rounding up).

	%	MIDTERM GRADE	FINAL GRADE	POINTS
100-98	A+			
97-96	A		A (passed -	
95-90	A-		very good)	450.
89-86	B+			
85-80	B		B (passed -	
79-76	B-		good)	380.
75-70	C+			
69-66	C		C (passed -	
65-60	C-		satisfactory)	300.
59-56	D+			
55-50	D		D (passed)	250.
49-0	F		F (failed)	

8. (Revision) Tests (ungraded)

- Complete before coming to class
- Recall last chapter or section of class
- Test question subset becomes final exam

9. Home assignments

- Complete DataCamp lessons
- Complete assignments outside of class
- Be ready to present your solutions

10. Class assignments

- Complete assignments in class (participation)
- Upload completed assignments (homework)
- Be ready to present your results in class

11. (Project) sprint review

- Research exploratory data analysis project
- Apply agile project management methods
- Present your results during 3 sprint reviews

12. Standard Policies

Honor Code

All graded work in this class is to be pledged in accordance with the Lyon College Honor Code. The use of a phone for any reason during the course of an exam is considered an honor code violation.

Class Attendance Policy

Students are expected to attend all class periods for the courses in which they are enrolled. They are responsible for conferring with individual professors regarding any missed assignments. Faculty members are to notify the Registrar when a student misses the equivalent of one, two, three, and four weeks of class periods in a single course. Under this policy, there is no distinction between “excused” and “unexcused” absences, except that a student may make up work missed during an excused absence. A reminder of the college’s attendance policy will be issued to the student at one week, a second reminder at two weeks, a warning at three weeks, and notification of administrative withdrawal and the assigning of an “F” grade at four weeks. Students who are administratively withdrawn from more than one course will be placed on probation or suspended.

Disabilities

The Morrow Academic Center (MAC) helps students who want to improve grades by providing peer-led services including Supplemental Instruction (SI), tutoring, the Writing Center, and academic coaching as well providing 24-hour, online tutoring for all subjects through Tutor.com. A schedule of peer-led services is available at lyon.edu/mac and Tutor.com is accessed through courses in Schoology. Contact Donald Taylor, Director of Academic Support, at 870-307-7319 or donald.taylor@lyon.edu for more information about MAC services.

Technology Support

For general technology support, you can contact the IT department by emailing support@lyon.edu or by calling 870-307-7555. For assistance with classroom-related technologies, such as the learning management system (LMS), you can request support using the methods above, or you can contact sarah.williams@lyon.edu directly for assistance. Your course content will be accessible digitally using either the Schoology or Canvas LMS. Both LMS platforms will use your myLyon

credentials for your student login.

- For Canvas, login at lyon.instructure.com
- For Schoology, login at lyon.schoology.com

Disabilities

Students seeking reasonable accommodations based on documented learning disabilities must contact Interim Director of Academic Support Donald Taylor in the Morrow Academic Center at (870) 307-7019 or at donald.taylor@lyon.edu.

Harassment, Discrimination, and Sexual Misconduct

Lyon College seeks to provide all members of the community with a safe and secure learning and work environment that is free of crime and/or policy violations motivated by discrimination, sexual and bias-related harassment, and other violations of rights. The College has a zero-tolerance policy against gender-based misconduct, sexual assault, and interpersonal violence toward any member or guest of the Lyon College community. Any individual who has been the victim of an act of violence or intimidation is urged to make an official report by contacting a campus Title IX coordinator or by visiting www.lyon.edu/file-a-title-ix-report. A report of an act of violence or intimidation will be dealt with promptly. Confidentiality will be maintained to the greatest extent possible within the constraints of the law. For more information regarding the College's Title IX policies and procedures, visit www.lyon.edu/title-ix.

Mental & Behavioral Health

Lyon College is dedicated to ensuring each student has access to mental and behavioral health resources. The College's Mental and Behavioral Health Office is located in Edwards Commons and is partnered with White River Health System's Behavioral Health Clinic. The office is committed to helping the Lyon community achieve maximum mental and behavioral wellness through both preventative and reactive care. A full-time, licensed, professional counselor provides counseling, consultations, outreach, workshops, and many more mental and behavioral services to Lyon students, faculty, and staff at no cost. The Mental and Behavioral Health Office also provides access to White River Health System's services and facilities, including medication management and in-patient and out-patient care. To make an appointment, contact counseling@lyon.edu.

College-Wide COVID-19 Policies for Fall, 2022

The College does not require masks in instructional and meeting spaces inside academic buildings. However, if instructors require masks in their classroom, lab, or studio, then students and guests must comply with that requirement. Vaccines are strongly encouraged for all faculty, staff, and students. Vaccines are not mandated for Lyon College community members, although there may be specific courses involving interactions with vulnerable, external populations where a vaccine may be required. The College will continue to offer symptomatic testing for students, faculty and staff.

Details

Details specific to this course may be found in the subsequent pages of this syllabus. Those details will include at least the following:

- A description of the course consistent with the Lyon College catalog.

- A list of student learning outcomes for the course.
- A summary of all course requirements.
- An explanation of the grading system to be used in the course.
- Any course-specific attendance policies that go beyond the College policy.
- Details about what constitutes acceptable and unacceptable student collaboration on graded work.
- A clear statement about which LMS is being used for the course.

13. Learning Management System (LMS)

We will use Canvas in this course (lyon.instructure.com/courses/655).

14. Assignments and Honor Code

There will be several assignments during the summer school, including programming assignments and multiple-choice tests. They are due at the beginning of the class period on the due date. Once class begins, the assignment will be considered one day late if it has not been turned in. Late programs will not be accepted without an extension. Extensions will **not** be granted for reasons such as:

- You could not get to a computer
- You could not get a computer to do what you wanted it to do
- The network was down
- The printer was out of paper or toner
- You erased your files, lost your homework, or misplaced your flash drive
- You had other coursework or family commitments that interfered with your work in this course

Put “Pledged” and a note of any collaboration in the comments of any program you turn in. Programming assignments are individual efforts, but you may seek assistance from another student or the course instructor. You may not copy someone else’s solution. If you are having trouble finishing an assignment, it is far better to do your own work and receive a low score than to go through an honor trial and suffer the penalties that may be involved.

What is cheating on an assignment? Here are a few examples:

- Having someone else write your assignment, in whole or in part
- Copying an assignment someone else wrote, in whole or in part
- Collaborating with someone else to the extent that your submissions are identifiably very similar, in whole or in part
- Turning in a submission with the wrong name on it

What is not cheating? Here are some examples:

- Talking to someone in general terms about concepts involved in an assignment
- Asking someone for help with a specific error message or bug in your program

- Getting help with the specifics of language syntax or citation style
- Utilizing information given to you by the instructor

Any assistance must be clearly explained in the comments at the beginning of your submission. If you have any questions about this, please ask or review the policies relating to the Honor Code.

Absences on Days of Exams: Test “make-ups” will only be allowed if arrangements have been made prior to the scheduled time. If you are sick the day of the test, please e-mail me or leave a message on my phone before the scheduled time, and we can make arrangements when you return.

15. Attendance policy

In accordance with college policy, if you miss 4 weeks of class, you fail the course automatically. Any missed meetings result in an ["Early Alert" report](#). You should take care not to miss consecutive sessions if at all possible - otherwise you risk losing touch with the class and falling behind.

16. Dates and class schedule

WHEN	WHAT
August 16	Classes begin
August 29	Last day to drop course
September 5	Labor day (no classes)
October 1-4	Fall break (no classes)
October 6	Mid-term grades due
October 14-26	Sesquicentennial week
October 20	Service day (no classes)
November 23-27	Thanksgiving break (no classes)
December 2	Last day of classes
December 14	Final grades due

WK	NO	DATE	TESTS	TEXTBOOK	ASSIGNMENT ¹
1	1	Tue-16-Aug	Entry Test		1. Emacs tutorial

¹ Assignments 2-15 are DataCamp assignments from the following courses: 1) [Introduction to statistics](#) (2-5), 2) [Introduction to statistics in R](#) (6-9), 3) [Foundations of probability in R](#) (10-13), and 4) [Introduction to regression in R](#) (14-15).

WK	NO	DATE	TESTS	TEXTBOOK	ASSIGNMENT
	2	Thu-18-Aug			
2	3	Tue-23-Aug	Test 1	13. Elementary statistics	2. Introduction to statistics
	4	Thu-25-Aug			
3	5	Tue-30-Aug	Test 2		3. Probability and distributions
	6	Thu-01-Sep			1st Sprint Review
4	7	Tue-06-Sep	Test 3		4. Central Limit Theorem
	8	Thu-08-Sep			
5	9	Tue-13-Sep	Test 4	14. Basic data visualization	5. Correlation/experiments
	10	Thu-15-Sep			
6	11	Tue-20-Sep	Test 5		6. Summary statistics in R
	12	Thu-22-Sep			
7	13	Tue-27-Sep	Test 6		7. Random numbers/probability in R
	14	Thu-29-Sep			2nd Sprint Review
8	15	Thu-06-Oct	Test 7	15. Probability	8. Central limit theorem in R
9	16	Tue-11-Oct	Test 8		9. Correlation/experiments in R
	17	Thu-13-Oct			
10	18	Tue-18-Oct	Test 9		10. The binomial distribution
11	19	Tue-25-Oct	Test 10		11. Laws of probability
	20	Thu-27-Oct			3rd Sprint Review

WK	NO	DATE	TESTS	TEXTBOOK	ASSIGNMENT
12	21	Tue-01-Nov	Test 11	16. Distributions	12. Bayesian statistics
	22	Thu-03-Nov			
13	23	Tue-08-Nov	Test 12		13. Related distributions
	24	Thu-10-Nov			
14	25	Tue-15-Nov	Test 13	20. Linear regression	14. Simple linear regression
	26	Thu-17-Nov			
15	27	Mon-22-Nov	Test 14		15. Predictions and model objects
	28	Tue-29-Nov			Final Sprint review
	29	Thu-01-Dec			Final Sprint review