UW-Madison Syllabus

General Course Information

Statistics 340 Data Science Modelling II

Credits: 4 credits
Level: Intermediate
L&S credit type: Counts as LAS credit (L&S)

Course Description

Teaches how to explore, model, and analyze data using R. Topics include basic probability models; the central limit theorem; Monte Carlo simulation; one- and two-sample hypothesis testing; Bayesian inference; linear and logistic regression; ANOVA; the bootstrap; random forests and cross-validation. Features the analysis of real-world data sets and the communication of findings in a clear and reproducible manner within a project setting.

Requisites

MATH 211, 217, 221, or 275. And STAT 240

Meeting Time and Location

Lecture 001 TR 1:00-2:15 204 Educational Sciences

Discussion:

311	W 4:35-5:25	375 Van Hise Hall	Nabil Awan
312	R 7:45-8:35	382 Van Hise Hall	Nabil Awan
313	W 3:30-4:20	B223 Van Vleck Hall	Nabil Awan
314	W 4:35-5:25	594 Van Hise Hall	Christian Varner
315	R 7:45-8:35	378 Van Hise Hall	Christian Varner
316	R 8:50-9:40	2255 Engineering Hall	Christian Varner

Instructional Modality

Lecture: In-person
Discussion: In-person
Office Hours: In person & Online – see canvas for schedule

How Credit Hours are Met by the Course

This course is 4 credits. The class meets for 150 minutes of lecture weekly and one 50-minute in-person discussion each week and carries the expectation that students will work on course learning activities (readings, homeworks, studying, etc.) for about 3 hours out of the classroom for every lecture period.

Other Course Information

Instructors & Teaching Assistants

Lecturer: Brian Powers
Office Hours: TBA on Canvas
Email: brpowers2@wisc.edu (for private communication)
(Please post all course-related questions to Discord so all students can benefit)

Teaching Assistants
311, 312, 313 Nabil Awan nawan2@wisc.edu
314, 315, 316 Christian Varner varner@wisc.edu
See Office Hours Grid on Canvas for most up to date office hours information

Course Learning Outcomes

Students will learn how to explore and analyze data using R, as well as how to present their findings and analyses clearly. Topics include basic probability models; the central limit theorem; Monte Carlo simulation; one- and two-sample hypothesis testing; Bayesian inference; linear and logistic regression; ANOVA; the bootstrap; cross-validation. Students will learn how to present their findings in a clear and reproducible manner in a project setting by applying their skills to analyze real-world data sets.

Grading

Students' final course score will be weighted according to the following percentages:

Component	Percentage
Discussion	10%
Homework	30%
Midterm Exam	30%
Final Exam	30%

Final Grading Scheme

Grade	A	AB	B	BC	C	D	F
Raw Score	[92, 100]	[88, 92)	[82, 88)	[78,82)	[70,78)	[60,70)	[0,60)

Course Website, Learning Management System & Digital Instructional Tools

Course Canvas Page: https://canvas.wisc.edu/courses/427090

Discussion Sessions

Discussion will be a chance to review lecture material, work through examples, and ask questions. Discussion will also include group work which will be submitted and graded on completion. Dicussion materials will be available to peruse the Friday before discussion. Students will be allowed to drop the lowest discussion score.

Textbook

There is no required textbook. Necessary reference materials will be posted to Canvas.

Software & Other Course Materials

We will be using R, and optionally, an R interface called R Studio. R is a free, open-source, and extremely flexible package, and is available for download online at: www.cran.r-project.org/ . R Studio is available for free from www.rstudio.com/products/rstudio/download/#download . Some experience with R will be expected, as R should be familiar from 240.

Homework

There will be approximately 11 (or so) weekly homework assignments throughout the semester. Generally, assignments will be posted to Canvas on Friday the week before they are due. Homework will be due the following Friday by midnight, with some exceptions (e.g. exam week). Submission will be online using Canvas. Extensions may be available in certain circumstances, but no homework will be accepted once the solutions have been posted (generally the Monday following the due date). Students will get to drop their lowest **two** homework scores. You must write up the homework solutions yourself to receive credit. Computer code and output must also be your own. Further details about guidelines and expectations for homework assignments will be provided on the assignments themselves.

Exams

There will be two midterm exams (one in-class and one at-home) and one in-class final exam.

- The in-class midterm exam will be in lecture on Thursday October 24
- The at-home midterm exam will be tentatively due Friday November 15 by midnight. Details will be posted on Canvas closer to this date.
- The final exam is Friday Dec 18, 12:15-2:15, location TBD.

Alternate dates for the in-class midterm exam will be observed only in extenuating circumstances. There will be no make-ups allowed for the final exam, for any reason, so please plan ahead. For the at-home midterm, you can use any resources you like, apart from other people.

Academic Integrity

Students are expected to abide by the rules of academic honesty. So, all graded assignments and exams must be compose of your own work. If there are any doubts or confusion, you must seek guidance from the instructor before submitting work.

Accomodations for Students with Disabilitites

The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. Students who need to arrange extra accomodations should contact both the instructor and the McBurney Disability Resource Center within the first two weeks of the term.

Teaching & Learning Data Transparency Statement

The privacy and security of faculty, staff and students' personal information is a top priority for UW-Madison. The university carefully *evaluates* and vets all campus-supported digital tools used to support teaching and learning, to help support success through *learning analytics*, and to enable proctoring capabilities. View the university's full *teaching and learning data transparency statement*.

Privacy of Student Records & the Use of Audio Recorded Lectures Statement

See more information about privacy of student records and the usage of audio-recorded lectures.

Lecture materials and recordings for this course are protected intellectual property at UW-Madison. Students in this course may use the materials and recordings for their personal use related to participation in this class. Students may also take notes solely for their personal use. If a lecture is not already recorded, you are not authorized to record my lectures without my permission unless you are considered by the university to be a qualified student with a disability requiring accommodation. [Regent Policy Document 4-1] Students may not copy or have lecture materials and recordings outside of class, including posting on internet sites or selling to commercial entities. Students are also prohibited from providing or selling their personal notes to anyone else or being paid for taking notes by any person or commercial firm without the instructor's express written permission. Unauthorized use of these copyrighted lecture materials and recordings constitutes copyright infringement and may be addressed under the university's policies, UWS Chapters 14 and 17, governing student academic and non-academic misconduct.

How to Succeed in This Course

The successful student will follow videos, lecture notes, and discussion materials, submit all homework and quizzes, take exams well-prepared, and utilize lecture, discussion, Piazza and office hours to ask questions when things are unclear.

Other Campus resources:

- *University Health Services*
- *Undergraduate Academic Advising and Career Services*
- *Office of the Registrar*
- *Office of Student Financial Aid*
- *Dean of Students Office*

Course Evaluations

Students will be provided with an opportunity to evaluate this course and your learning experience. Student participation is an integral component of this course, and your confidential feedback is important to me. I strongly encourage you to participate in the course evaluation.

Digital Course Evaluation (AEFIS)

UW-Madison uses a digital course evaluation survey tool called AEFIS. For this course, you will receive an official email two weeks prior to the end of the semester, notifying you that your course evaluation is available. In the email you will receive a link to log into the course evaluation with your NetID. Evaluations are anonymous. Your participation is an integral component of this course, and your feedback is important to me. I strongly encourage you to participate in the course evaluation.

Students' Rules, Rights & Responsibilities

Diversity & Inclusion Statement

Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals. The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world.

Academic Integrity Statement & Data Ethics

By virtue of enrollment, each student agrees to uphold the high academic standards of the University of Wisconsin-Madison; academic misconduct is behavior that negatively impacts the integrity of the institution. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these previously listed acts are examples of misconduct which may result in disciplinary action. Examples of disciplinary action include, but is not limited to, failure on the assignment/course, written reprimand, disciplinary probation, suspension, or expulsion.

The members of the faculty of the Department of Statistics at UW-Madison uphold the highest ethical standards of teaching, data, and research. They expect their students to uphold the same standards of ethical conduct. Standards of ethical conduct in data analysis and data privacy are detailed on the ASA website (https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice), and include:

- Use methodology and data that are relevant and appropriate; without favoritism or prejudice; and in a manner intended to produce valid, interpretable, and reproducible results.
- Be candid about any known or suspected limitations, defects, or biases in the data that may affect the integrity or reliability of the analysis. Obviously, never modify or falsify data.
- Protect the privacy and confidentiality of research subjects and data concerning them, whether obtained from the subjects directly, other persons, or existing records.

By registering for this course, you are implicitly agreeing to conduct yourself with the utmost integrity throughout the semester.

Netiquette on Discord and online communication

See https://kb.wisc.edu/50548 for a general netiquette. Specifically:

- Any comment or answers must be on topic, concise, polite, and respectful of others.
- Assume the best intentions of others in the class and be forgiving when you think that the tone of someone's post is offensive. It is easy to misread the tone of someone's written communication. If in doubt, ask an open, honest question about what the person meant so that you can clarify before making assumptions that damage your perception of your colleague.
- Students must not post answers to homework problems.
- Almost all questions should be sent via Piazza. For personal or sensitive issues, use private messages to instructor/TA in Piazza, or email.
- Students should not expect an immediate answer to a question posted late at night before an assignment due date.

Statement on the use of ChatGPT and other AI language models

While the Statistics Department recognizes the potential benefits of AI, its use in academic work can be problematic. In this course, two rules regarding the use of ChatGPT and other generative AI models will be enforced: (1) Passing off AI-generated responses as original student work constitutes plagiarism and is strictly prohibited. Any students found to be engaging in this practice will be cited for academic misconduct. (2) Unless explicitly authorized by the instructor to do so, any use of AI-generated responses as sources of information, even with documentation and attribution, is prohibited.

Complaints

If you have a complaint about a TA or course instructor, you should feel free to discuss the matter directly with the TA or instructor. If the complaint is about the TA and you do not feel comfortable discussing it with him or her, you should discuss it with the course instructor. Complaints about mistakes in grading should be resolved with the instructor in the great majority of cases. If the complaint is about the instructor (other than ordinary grading questions) and you do not feel comfortable discussing it with him or her, contact the Director for Undergraduate Studies, Professor Jessi Cisewski-Kehe (jjkehe@wisc.edu).

If your complaint concerns sexual harassment, please see campus resources listed at https://compliance.wisc.edu/titleix/. In particular, there are a number of options to speak to someone confidentially.

If you have concerns about climate or bias in this class, or if you wish to report an incident of bias or hate that has occurred in class, you may contact the Chair of the Statistics Department Climate & Diversity Committee, Professor Karl Rohe (karl.rohe@wisc.edu). You may also use the University's bias incident reporting system, which you can reach at https://osas.wisc.edu/report-an-issue/bias-or-hate-reporting/.

Overlapping course time statement

The Department of Statistics strongly discourages students from enrolling in any courses whose regular class meeting dates & times overlap with each other. This policy is in alignment with the College of Letters and Sciences Course Attendance Policy. It is also consistent with the Class Attendance Policy for Students at UW-Madison (https://kb.wisc.edu/ls/24628), whose first sentence reads, "It is expected that every student will be present at all classes." Statistics instructors may opt not to make any alternative arrangements in the event any conflict arises due to a student taking a course with class meetings that overlap with a Statistics course, including a conflict between two Statistics courses. Note that final exams occasionally are scheduled simultaneously for courses which meet at different times; in this situation, please contact your instructor well before the exam date about potential accommodations.

Accommodations for Students with Disabilities Statement

The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (UW-855) require the university to provide reasonable accommodations to students with disabilities to access and participate in its academic programs and educational services. Faculty and students share responsibility in the accommodation process. Students are expected to inform faculty [me] of their need for instructional accommodations during the beginning of the semester, or as soon as possible after being approved for accommodations. Faculty [I], will work either directly with the student [you] or in coordination with the McBurney Center to provide reasonable instructional and course-related

accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA. (See: McBurney Disability Resource Center)

Academic Calendar & Religious Observances

Plase contact Dr. Powers as soon as you are able about any scheduling changes that need to happen to accommodate your religious observances.