

CS 541: Advanced Data Management

Lecture: Tuesday 12:10pm - 3:10pm

Address: Online

Zoom Link: Meeting ID: 937 3245 6014 Password: 476278

<https://rutgers.zoom.us/j/93732456014?pwd=NHl5bXFIL1hKRohvUGc4NTIzMHhOZz09>

Pre-requisites:

(01:198:214 SYSTEMS PROGRAMMING) AND (01:198:336 PRIN INFO & DATA MGT)

Instructor: Dong Deng (dong.deng AT rutgers.edu)

Office Hour: By appointment

Course Description

This course is designed to introduce some advanced topics in data management and data curation, including but not limited to: data integration, data cleaning, data wrangling, and data visualization. Classes consist of lectures and seminars. There will be two programming assignments, one exam, and one project presentation.

Lectures are held once a week. Attendance at lectures is mandatory and you are expected to show up and participate in discussion.

Grading

The final grade for the course will be based on the following weights:

- 40% — Programming Assignments
- 30% — Project and Presentation
- 30% — Open book Examination
- 5% — Discussion, Participation, etc. (at the lecture's discretion)

The final grade depends on the percentage of points you have earned, and the definition of letter grades is:

- A: 85 – 100
- B+: 75 – 84
- B: 65 – 74
- C+: 55 – 64
- C: 45 – 54
- F: < 44

Reading Materials

all lecture discussions are based on readings from the data management literature.

Academic Integrity

Students are expected to maintain the highest level of academic integrity. You should be familiar with the university policy on academic integrity. Violations will be reported and enforced according to this policy. Use of external website resources such as Chegg.com or others to obtain solutions to homework assignments, quizzes, or exams is cheating and a violation of the University Academic Integrity policy. Cheating in the course may result in grade penalties, disciplinary sanctions or educational sanctions. Posting homework assignments, or exams, to external sites without the instructor's permission may be a violation of copyright and may constitute the facilitation of dishonesty, which may result in the same penalties as plain cheating.

Rutgers CS Diversity and Inclusion Statement

Rutgers Computer Science Department is committed to creating a consciously anti-racist, inclusive community that welcomes diversity in various dimensions (e.g., race, national origin, gender, sexuality, disability status, class, or religious beliefs). We will not tolerate micro-aggressions and discrimination that creates a hostile atmosphere in the class and/or threatens the well-being of our students. We will continuously strive to create a safe learning environment that allows for the open exchange of ideas and cherished freedom of speech, while also ensuring equitable opportunities and respect for all of us. Our goal is to maintain an environment where students, staff, and faculty can contribute without the fear of ridicule or intolerant or offensive language. If you witness or experience racism, discrimination micro-aggressions, or other offensive behavior, you are encouraged to bring it to the attention to the undergraduate program director and/or the department chair. You can also report it to the Bias Incident Reporting System <http://inclusion.rutgers.edu/report-bias-incident/>

Student-Wellness Services

Counseling, ADAP & Psychiatric Services (CAPS)

(848) 932-7884 / 17 Senior Street, New Brunswick, NJ 08901 <http://health.rutgers.edu/medical-counseling-services/counseling/>

CAPS is a University mental health support service that includes counseling, alcohol and other drug assistance, and psychiatric services staffed by a team of professionals within Rutgers Health services to support students' efforts to succeed at Rutgers University. CAPS offers a variety of services that include: individual therapy, group therapy and workshops, crisis intervention, referral to specialists in the community, and consultation and collaboration with campus partners.

Crisis Intervention : <http://health.rutgers.edu/medical-counseling-services/counseling/crisis-intervention/>

Report a Concern: <http://health.rutgers.edu/do-something-to-help/>

Violence Prevention & Victim Assistance (VPVA)

(848) 932-1181 3 Bartlett Street, New Brunswick, NJ 08901 www.vpva.rutgers.edu/

The Office for Violence Prevention and Victim Assistance provides confidential crisis intervention, counseling and advocacy for victims of sexual and relationship violence and stalking to students, staff and faculty. To reach staff during office hours when the university is open or to reach an advocate after hours, call 848-932-1181.

- Disability Services

(848) 445-6800 *Lucy Stone Hall, Suite A145, Livingston Campus, 54 Joyce Kilmer Avenue, Piscataway, NJ 08854* <https://ods.rutgers.edu/>

Rutgers University welcomes students with disabilities into all of the University's educational programs. In order to receive consideration for reasonable accommodations, a student with a disability must contact the appropriate disability services office at the campus where you are officially enrolled, participate in an intake interview, and provide documentation: <https://ods.rutgers.edu/students/documentation-guidelines>. If the documentation supports your request for reasonable accommodations, your campus's disability services office will provide you with a Letter of Accommodations. Please share this letter with your instructors and discuss the accommodations with them as early in your courses as possible. To begin this process, please complete the Registration form on the ODS web site at: <https://ods.rutgers.edu/students/registration-form>.

Tentative Schedule

	Date	Note	Topic
1	1/18		Introduction and Preliminary
2	1/25		String Matching – Partition-based Algorithms
3	2/1		String Matching – Heap-based Algorithms
4	2/8		String Matching – Prefix Filtering
5	2/15		Nearest Neighbor Search in Multi-Dimension
6	2/22		Nearest Neighbor Search in High-Dimension
7	3/1		Program Synthesis for Data Wrangling
8	3/8		All-Pair Near-Duplicate Text Detection
	-		
9	3/22		Exam
10	3/29		Project Presentation – 1
11	4/5		Project Presentation – 2
12	4/12		Project Presentation – 3
13	4/19		Project Presentation – 4
14	4/26		Project Presentation – 5