

UNIVERSITY OF CALIFORNIA, DAVIS
GRADUATE SCHOOL OF MANAGEMENT

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO



SANTA BARBARA • SANTA CRUZ

BAX 422 - Data Design & Representation (2 units)

Winter 2025

Preliminary Syllabus

Instructor: Dr. Sabid Rahman

E-Mail: krahman@ucdavis.edu

Office Hours:

Every Monday 5PM - 6PM. Virtual meeting links will be shared.

Please feel free to reach out to me via email at any point. We can also schedule additional meetings if necessary.

Objectives:

This course will focus on automatically extracting information from unstructured sources, process it, and structure it for business usecases.

Key topics include:

Regular Expressions (Regex)

Web-Scraping

Web Data Processing

JavaScript Object Notation (JSON)

MongoDB (NoSQL)

Textbook:

There is **no required textbook**. Topics of this course are diverse. We may review materials from specific books, but all related materials will be available on the class canvas. The best preparation you can do before lectures is to go through class code and work through examples.

Teaching Assistant: TBD

TA office hours: TBD. Virtual.

Grading:

Assignments: 20%

Project 1 : 20%

Project 2: 20%

Midterm: 20%

Final: 20%

Midterm: Feb. 15 (in person in class).

Final exam: March 15 (online exam).

Class location details:

In person bi-weekly class

Location:

200 McAllister St., Room 123 (1st Floor),
San Francisco, CA 94102

Refer to the SAN FRANCISCO CAMPUS GUIDE uploaded in Canvas for parking and directions.

Tentative timelines:

Date (Weeks)	Topics
01/04/2025 (1&2)	Intro, Web scraping and Regular Expressions
01/18/2025 (3&4)	Web scraping and Web data processing
02/01/2025 (5&6)	Web data processing and JSON
02/15/2025 (7&8)	Web data processing and MongoDB (Midterm)
03/01/2025 (9&10)	MongoDB and Review

Course guidelines and information:

- Class participation is key to the success of the course learning experience.
- Use of AI is prohibited by default UC Davis guideline.
- Late assignments, exams, or projects, will not be accepted.
- Clerical scoring errors will be corrected without hassle, but for other re-grades you must hand back the work and submit an email request; the entire paper will be subject to regrading.
- Projects will be group assignments. You can pick your group of 3-4 persons. If you do not have one, we will help you to assign a team.
- Projects will be assigned in week 4 and 8. They will be due in 10 days.
- Successful projects will require programming, database and managerial skills.
- Python will be used for the programming assignments and projects in this course.
- This course is not a programming course. But a basic understanding of Python will be helpful to easily follow the code developed in class.

Academic Code of Conduct:

- All participants in the course, instructor and students, are expected to follow the [UC Davis Principles of Community](#).
- You are expected to take UC Davis's [Code of Academic Conduct](#) as seriously as we do.

Reference materials (for students' further support; not textbooks):

- Introduction to [Python Programming in Python](#).
- Mastering Regular Expressions by Jeffrey Friedl.
- Ryan Mitchell's Web Scraping with Python: Collecting More Data from the Modern Web
- Data Modeling Essentials by Simson Witt