INST 447 Data Source and Manipulation

Term:	Fall 2025
Instructor:	Wei Ai (aiwei@umd.edu)
Lecture:	Thursday. 12:30 - 1:45 pm. TMH 0301
Lab:	Tuesday. 12:30 - 1:45; 2:00 - 3:15; 3:30 - 4:45; HBK 0302J
Office Hours:	Wednesday. 11 - 12 (virtual) 4 - 5 (in-person, room TBD)
	More office hours to be announced.

Course Description

This course examines approaches to locating, acquiring, manipulating, and disseminating data. We study practical methods for gathering data (e.g., APIs, basic web data), transforming common formats (CSV, JSON, XML), cleaning and validating tables, and wrangling tasks such as aggregation, subsetting, merging, and reshaping. Consistent with current professional practice, students are encouraged to try AI coding assistants and LLM tools for coding support and data tasks.

Learning Outcomes

By the end of the course, students will be able to:

- **Locate, acquire, and organize data** from files, APIs, and simple web sources.
- **Clean and standardize datasets** to prepare them for analysis.
- **Transform and reshape data** across common formats and structures.
- **Combine datasets and summarize information** to answer practical questions.
- **Document decisions and validate results** with quick checks
- **Communicate methods and findings clearly** for non-technical audiences.
- **Use AI assistants responsibly** to draft code and automate routine data tasks, while specifying constraints and verifying outputs (schema checks, key-uniqueness tests, row-delta checks). (New emphasis; aligned with course goals above.)

Required Resources

This course does not require any textbook. Pointers to a selection of readings will be provided via ELMS.

Course Structure

Lectures: Conceptual foundations and live demonstrations of data acquisition, cleaning, transformation, and validation, with examples of specifying tasks precisely and verifying outcomes.

Labs: Guided, hands-on practice applying the previous lecture's concepts. Each lab closes with a brief reflection on what you validated and—when applicable—how you used AI.

Programming Exercises: Short, more rigorous checkups that consolidate recent skills (e.g., filtering+types; group & reshape; JSON/API; text/regex). You submit a working notebook plus a concise process/validation note.

Take-Home Midterm: Practical, windowed assessment integrating the first-half skills into a small, realistic data task.

Final Project: A compact, end-to-end analysis where you source/prepare data, explore it, and produce a stakeholder-oriented write-up; responsible use of AI tools is welcome with appropriate validation.

Grading

Grading Structure

- Weekly Labs: 30%
- In-Lecture Quizzes: 10%
- Programming Assignments (4): 20%
- Take-Home Midterm: 20%
- Final Project: 20%

Letter Grade Cutoffs

A+	97-100*	B+	87-89.99	C+	77-79.99	D+	67-69.99	F	0-59.99
A	93-96.99	B	83-86.99	C	73-76.99	D	63-66.99		
A-	90-92.99	B-	80-82.99	C-	70-72.99	D-	60-62.99		

*Note: To receive an A+ you must have demonstrated significant contributions to the class in addition to achieving this numeric grade. We reserve the right to curve grades upward (but will not curve grades downward).

Communications

- ELMS: Official course site for materials, assignments, announcements, gradings, etc.
- Emails: Administrative requests, etc. ***Please prefix the subject line with [INST750].*** If you have not received a reply within 2 days, please email again.
- Office Hour: preferred for in-depth and/or back-and-forth discussions.
- **Please do not wait until last-minute to communicate!**

Tentative Schedule

Class #	Date	Topic
1	09/04	Introduction, course workflow & pipeline
2	09/11	Filtering & transformation
3	09/18	Grouping, joining & reshaping
4	09/25	Visualization for validation; thinking at scale (map-reduce)
5	10/02	Working with text: string methods & practical regex
6	10/09	Semi-structured data: JSON/XML
7	10/16	APIs & web data
8	10/23	AI for data processing
9	10/30	Data representation abstractions
10	11/06	Practical vibe coding: iterative exploration with AI
11	11/13	The data quality pipeline: missing data, outliers, dedupe, etc.
12	11/20	Responsible analysis: brief bias/ethics audit & documentation
	11/27	(Thanksgiving break)
13	12/04	Advanced topics TBD
14	12/11	Course synthesis & reflection

Campus Policies

It is our shared responsibility to know and abide by the University of Maryland's policies that relate to all courses, which include topics like:

- Academic integrity
- Student and instructor conduct
- Accessibility and accommodations
- Attendance and excused absences
- Grades and appeals
- Copyright and intellectual property

Please visit go.umd.edu/ug-policy for the Office of Undergraduate Studies' full list of campus-wide policies and follow up with me if you have questions.

Additional Accommodation Policy

I understand the difficulty and additional constraints you may be facing during this time. I am willing to work with you to discuss possible accommodation and alternative arrangement. Please don't hesitate to contact me when needed.

Accessibility and Learning Support

Students with disabilities should inform me of their needs at the beginning of the semester. Please also contact the Accessibility and Disability Support Office (http://www.counseling.umd.edu/ADS/). ADS will make arrangements with the student and me to determine and implement appropriate academic accommodations. Inclusion is one of the iSchool's core values, and we have attempted to make all materials and assignments accessible to people with varying abilities. However, if there is something else I can do to make the class more accessible please schedule a time to come talk to me. This will benefit not only yourself but also future students!

Get Some Help!

Taking personal responsibility for your own learning means acknowledging when your performance does not match your goals and doing something about it. I hope you will come talk to me so that I can help you find the right approach to success in this course, and I encourage you to visit tutoring.umd.edu to learn more about the wide range of campus resources available to you. In particular, everyone can use some help sharpen their communication skills (and improving their grade) by visiting ter.ps/writing and schedule an appointment with the campus Writing Center. You should also know there are a wide range of resources to support you with whatever you might need (see go.umd.edu/assistance), and if you just need someone to talk to, visit counseling.umd.edu or one of the many other resources on campus.

Most services free because you have already paid for it, and everyone needs help… all you have to do is ask for it.