

ECON 8320 – Tools for Data Analysis

Updated Summer 2021

Instructor: Dustin White
MH 332M
Email: drwhite@unomaha.edu

Office Hours: 4:30-5:30 PM prior to class, and by appointment.

Materials: Required materials in **bold**

Course Slides (hosted on Github)
Course Notes (also hosted on Github)
Mimir Account (\$25 per semester)
Python Lectures PDF from Quant-Econ.net
Data Science from Scratch (ISBN: 978-1-4919-0142-7)

Prerequisites: ECON 2200 or BSAD 8150 (or equivalent); BSAD 2130 or equivalent.
No previous programming experience is required.

Description: The course will cover basic principles of programming languages, as well as libraries useful in collecting, cleaning and analyzing data in order to answer research questions. The course will utilize basic Economic principles and Econometric methods as inspiration for assignments and projects throughout the duration of the course, and will do so in a way that is accessible to non-Economists. This course is intended to introduce the student to the Python programming language as a tool for conducting data analysis. While the course uses Python, the student should be able to move to other languages frequently used in data analysis using the principles taught in this course.

Course Outline: Data Types and Documentation 1 day
Functions 1 day
Classes 1 day

Factoring and Debugging	1 day
Regex	1 day
Web Scraping (Scrapy)	1 day
Pandas (pandasql and sqlite3)	1 day
Numpy, Scipy	1 day
Scipy.optimize	1 day
Plotting (Plotly)	1 day
Natural Language Processing	1 day
Statsmodels, sklearn	1 day
Web APIs	1 day
Multithreading	1 day
 Project Presentations	 2 days

Grade Policy: Homework (1 assignment per topic)	60% of final grade
Semester Project	30% of final grade
Participation	10% of final grade

Each homework assignment is identically weighted in calculating final grades, but may be measured a different scale depending on the requirements of the assignment.

NO LATE WORK WILL BE ACCEPTED.

Grades will be distributed according to the following grade scale:

Score	Letter Grade	Score	Letter Grade
A	> 93.9%	C+	77.5% - 79.9%
A-	90.0% - 93.9%	C	72.5% - 77.4%
B+	87.5% - 89.9%	C-	70.0% - 72.4%
B	82.5% - 87.4%	D	60.0% - 69.9%
B-	80.0% - 82.4%	F	< 60.0%

Course Objectives: After this course, students should be capable of:

1. Programmatically collecting data from websites and other sources for analysis
2. Automating the data cleaning process, by preparing collected data for analysis
3. Using code to analyze data in order to draw conclusions about the real world from which decisions can be made

Participation: Students are expected to be actively engaged during class in discussions during lecture as well as during lab time, and to register for the course Slack Channel (link in Canvas). We learn from each other, and I expect each student to contribute to our learning environment. Excess absences or lack of participation will result in loss of participation points at the instructor's discretion.

Grading: All homework assignments are to be submitted through Mimir, in order for students to receive automated and instant feedback wherever

possible. This will allow students to iteratively improve their work and practice coding. Rubrics will be posted, and will contain detailed information on the assignment grading policy.

Projects will be submitted through the appropriate drop box on the course Canvas page.

Homework: In order to give students as many opportunities as possible to practice the concepts being taught in class, there will be homework assigned for each class period covering new content. I will not accept lab assignments from students not present in lab, in order to emphasize the importance of attending class each week. You will be given time to work with your peers and learn from each other as you practice coding. No late homework will be accepted.

Projects: There will be one term project, which will be discussed during class. This project will be designed to allow students to showcase an ability to synthesize and combine the tools introduced in class to create a workflow leading from data collection through data cleaning to a final data analysis.

Work Load: The best way to learn is to do, and so we will focus on actively using the tools we discuss in class. I don't expect you to know how to code when the semester starts, but the course will be based on writing code, so I do expect you to learn as the course progresses. I will help you do so, and will make the process as painless as possible. The primary goal is to help you do data analysis. Your entire grade is based on coding projects and assignments, so please make sure that you schedule time to remain for all of class each week.

Academic Honesty: UNO's requirements for Academic Integrity and Behavior All students are required to adhere to the highest standards of academic integrity and behavior and must satisfy the UNO Academic Integrity Policy <http://www.unomaha.edu/student-life/student-conduct-and-community-standards/policies/academic-integrity.php> and Student Code of

Conduct <http://www.unomaha.edu/student-life/student-conduct-and-community-standards/policies/code-of-conduct.php>. It is the students responsibility to read, understand and abide by these policies. If I find that you have plagiarized, been dishonest in completing your assignments, or cheated an an exam or assignment, then I reserve the right to award you no points on the entire exam, project, or assignment and to report the behavior to the university. If this behavior is repeated, I reserve the right to award a failing grade, independent of your score on other assignments. Academic integrity is essential to education, and I take it very seriously.

Extra Help: Dot not hesitate to come to my office during office hours or by appointment to discuss a homework problem or any aspect of the course. Remember to ask questions on Slack! I'll be there to answer questions!