DS2000 -- Intermediate Programming with Data
Khoury College, Northeastern University
Spring 2022
Profs. John Rachlin (he/him) and Laney Strange (she/her)

Instructor emails	j.rachlin@northeastern.edu laneys@northeastern.edu
Course web page	http://course.ccs.neu.edu/ds2000
Piazza	https://piazza.com/northeastern/spring2022/ds2000
Gradescope	https://www.gradescope.com/courses/334317 access code JBW3BP
Lecture Schedule	Sec 1 TF 9:50-10:55am. WVH 102. Sec 2: TF 1:35-2:40pm. WVG 102. Sec 3: TF 3:25-4:30pm. WVG 104. Sec 4: T 11:45am-1:25pm, R 2:50-4:30pm. Online only.
Practicum Schedule (DS2001)	Practicum 1. T 1:35-3:15pm Practicum 2. T 3:25-5:05pm Practicum 3. W 11:45am-1:25pm Practicum 4. W 11:45am-1:25pm Practicum 5. W 2:50-4:30pm Practicum 6. W 2:50-4:30pm Practicum 7. W 4:40-6:20pm. Practicum 8. R 9:50-11:30am Practicum 9. R 11:45am-1:25pm Practicum 10. R 11:45am-1:25pm.

DS2001 is a co-requisite for DS2000; make sure you're signed up for both.

Office Hours

Instructor Office Hours (Online)

Instructor office hours are one-on-one conversations. You can schedule a specific 20-minute slot during office hours. Priority is given to students with an appointment, but you can also just drop in.

- John Rachlin, TBD
    - Online: https://northeastern.zoom.us/my/rachlin
    - Make an appointment: https://calendly.com/rachlin/help
- Laney Strange, MW 10am-12pm
    - Online: https://northeastern.zoom.us/my/laney
    - Make an appointment: https://calendly.com/laneystrange/20-minute-office-hours-check-in

TA Office Hours (Online)

- Course TAs will hold regular office hours throughout the week. The website lists the schedule, and we'll post any changes on Piazza.
- Like an in-person office hour, you'll be joining a group of your classmates, with an TA or two hosting the meeting.
- If you need to check-in with someone one-on-one, please set up a time with the instructor.

Recommended Textbooks

- Intro to Python for Computer Science and Data Science. Deitel & Deitel. Pearson, 2019. ISBN: 0135404673. Available free online or purchase.
- Think Python: How to Think Like a Computer Scientist. Allen B. Downey. O'Reilly Media, 2015. ISBN: 1491939362. Available free online or purchase.

Relevant chapters will be listed alongside lecture topics on the course website. You do not need to read the textbooks ahead of lecture; they are most useful as reference materials or for looking up new examples. Keep them handy when working on the homework or reviewing your lecture notes.

Course Description - DS2000

Introduces programming for data and information science through case studies in business, sports, education, social science, economics, and the natural world. Presents key concepts in programming, data structures, and data analysis through Python. Integrates the use of data analytics libraries and tools. Surveys techniques for acquiring and programmatically integrating data from different sources. Explains the data analytics pipeline and how to apply programming at each stage. Discusses the programmatic retrieval of data from application programming interfaces (APIs) and from databases. Applies data visualization techniques to summarize and communicate the analysis of data.

Beginning programmers are welcome; we don't assume any previous knowledge and we'll start from the

very beginning.

The major topics within the course, and their corresponding textbook chapters, are the following (note that the order in which topics are covered might change):

Text Section(s) - Downey	**Topics**
Ch 1, 2.1-2.3	Variables, mathematical operators
Ch 9.1-9.3	Files and data visualization
Ch 3.1-3.7	Conditionals
Ch 5.17	Iteration (loops) and lists
Ch 5.7-5.8, 4	Functions
Ch 5.16-5.18	2D lists
Ch 6.1-6.2	Dictionaries
Ch 10	Classes & Objects
Ch 8.13-7.14	Pandas

Evaluation

Evaluation

You will receive separate grades for DS2000 and DS2001. Your DS2000 grade will be based on the following factors:

Factor	**Number**	**Weight**
Homework Sets HW1-7 are assigned work HW8 is a re-submission	7	75%
Quizzes (lowest dropped)	8	20%
Mini-Presentation	1	5%
	TOTAL	**100%**

You'll submit homeworks and quizzes through Gradescope. Allow at least one week after you submit a homework or exam before the grades are posted.

Letter Grades

Your final grade for DS2000 will use the following breakpoints to convert from letter to number grades.

Letter	Range
A	95-100
A-	90-94
B+	87-89
B	83-86
B-	80-82
C+	77-79
C	73-76
C-	70-72
D	60-69
F	< 60

Homeworks

Homeworks are assigned (almost) every week. They are due one week after they are assigned, unless otherwise noted.

Homeworks will be evaluated according to the DS2000 Grading Rubric. The final assignment of the semester, Homework 8, will be a make-up homework. You can use this homework to re-submit one of your earlier assignments, and we'll re-grade it. It's a chance to re-do a homework that didn't go as well as you'd hoped, or submit one where you'd missed the original deadline.

Your homework score will be the average of HW1-7.

Quizzes

We'll have weekly quizzes on Gradescope, given every Tuesday. The quizzes are auto-graded and you can tell if you got a problem wrong. You can resubmit your answers right up until the quiz deadline.

Mini-Presentation

Data scientists are great communicators. This one-on-one presentation is a chance to practice and get feedback on these non-Python aspects of data science: creating visualizations and communicating about them.

You'll generate a plot based on data we used in class, in your practicum, or on a homework. Your plot has to be new in some way -- something we didn't do as a group, and something that wasn't specifically assigned in a homework.

You'll present your plot in a remote one-hour time slot shared among 10 students. These presentations will take place in late February and early March. You'll be graded on the quality of your plot and the clarity of your presentation.

Late/Makeup Policy

- ***Homework*** - You can submit homeworks up to 24 hours late with a 5% penalty, and up to 48 hours late with a 10% penalty. No other late submissions will be accepted. This policy exists for those times you're having a tough week, are feeling sick, or are falling behind in your work; we won't make any exceptions to this policy.

- ***Quizzes*** - No late quizzes are accepted, but we will drop your lowest one. Quizzes are open from the end of class Tuesday (4:30pm) until the start of class Friday (9:50am). Gradescope tells you when you get something right, and you can resubmit right up until the deadline.

- ***Mini-Presentation*** - We expect you to be present in the time slot you signed up for, and we expect you to submit your plot by the deadline. In case of extenuating circumstances, reach out to the instructors at least 24 hours before your scheduled presentation.

Software

We'll be using Python 3 in this class. Anaconda (https://www.anaconda.com/) is your best bet for installing the latest version of Python along with various libraries.

When you install Anaconda, it also comes with the editor Spyder, which we'll use to write and run Python code. Spyder will be our "official" DS2500 editor; if you like and use another editor that's totally fine, but we'll use Spyder in lectures and office hours, and we'll be able to help you out if something goes wrong.

In addition to Spyder, we'll use Jupyter Notebooks in class and labs. Jupyter Notebooks are a great tool to have, but it is not required for any homework.

Communication

The simplest way to get feedback and help from course staff and from your classmates is via Piazza. Piazza is an extension of our classroom discussion, and we expect everyone to behave accordingly. No disrespect, rudeness, or abuse will be tolerated -- towards fellow students or towards the course staff. Piazza will be disabled if we feel it is being misused.

You may not post your code on Piazza, but you can ask, answer, and discuss different things you've tried, what worked and didn't work, and resources you've found.

We'll also use Piazza to post course announcements, so make sure your email settings are turned on!

Email (laneys@northeastern.edu, j.rachlin@northeastern.edu) is the best tool for specific questions or concerns about your experience in class or anything sensitive in nature. During the week, we'll respond within 24 hours, but don't expect a response after 9pm. On the weekends we'll be slower to respond, but if you reach out over a weekend you can expect to hear back by Sunday evening.

Office hours are the best place for talking through your approach to a homework problem. We're not here to give you answers, of course, but to be your fellow data scientists thinking through a tough problem with you. Expect us to ask more questions than we answer.

Classroom Environment

In our classroom, please ask questions, and answer questions! In programming, we seldom get anything right on the first try. We see how an attempt turned out, and we try again. I like our classroom to reflect that approach as well; so please answer a question that's been posed, even if you're not sure of the answer.

To create and preserve a classroom atmosphere that optimizes teaching and learning, all participants share a responsibility in creating a civil and non-disruptive forum for the discussion of ideas.

Students are expected to conduct themselves at all times in a manner that does not disrupt teaching or learning. Your comments to others should be constructive and free from harassing statements.

Academic Integrity

You are free to discuss homeworks and share code with your classmates. **You may not post code on piazza**.

Searching online and looking for ideas is acceptable, as long as (1) you cite any outside sources that you referenced in a comment in your code, and (2) you do not ask TAs or instructors to help you fix code you found online. We'll help you work out problems with your code, not someone else's.

Student Accessibility

If you require support during the course due to a disability please ensure that you are already registered with the University's Disability Center, and contact your course instructors to coordinate any support needed during the course.

Title IX makes it clear that violence and harassment based on sex and gender are Civil Rights offenses subject to the same kinds of accountability and the same kinds of support applied to offenses against other protected categories such as race, national origin, etc. If you or someone you know has been harassed or assaulted, you can find the appropriate resources here: Title IX.