DS2000 -- Intermediate Programming with Data
Khoury College, Northeastern University
Spring 2023
Prof. Laney Strange (she/her) and coordinator Kayla McLaughlin (she/her)

Emails	laneys@northeastern.edu k.mclaughlin@northeastern.edu
Course web page	http://course.ccs.neu.edu/ds2000
Piazza	https://piazza.com/northeastern/spring2023/ds2000
Gradescope	https://www.gradescope.com/courses/477682 access code 5754J3
Lecture Schedule	Sec 1 TF 9:50-10:55am. MU 201. Sec 2: TF 11:45am-12:50pm. EV 024. Sec 3: TF 3:25-4:30pm. EV 024.
Practicum Schedule (DS2001)	Practicum 1. W 9:50-11:30am Practicum 2. W 11:45am-1:25pm Practicum 3. W 2:50-4:30pm Practicum 4. W 2:50-4:30pm Practicum 5. W 4:40-6:20pm Practicum 6. R 9:50-11:30am Practicum 7. R 9:50-11:30am Practicum 8. R 11:45am-1:25pm Practicum 20. R 2:50-4:30pm Practicum 21. R 2:50-4:30pm Practicum 22. W 2:50-4:30pm Practicum 23. W 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 and In-Person)

Instructor office hours are one-on-one conversations. You can schedule a specific 15-minute slot during office hours. Priority is given to students with an appointment, but you can also just drop in. Both Zoom and in-person are perfectly fine with me, just let me know which you prefer when you make an appointment!

- Laney's office hours: M 2-4pm, R 9-11am
    - In-person: Meserve 313
    - Online: https://northeastern.zoom.us/my/laney
    - Make an appointment: https://calendly.com/laneystrange/20-minute-office-hours-with-laney

TA Office Hours (Online and In-Person)

- Course TAs will hold regular office hours throughout the week. The website lists the schedule, and we'll post any changes on Piazza.
- For in-person office hours, please go to the room listed and write your name at the bottom of the list on the whiteboard. A TA will call your name when it's your turn, so please keep an ear out!
- For online office hours, we will be using Zoom. Have a question ready and join the Zoom waiting room. When a TA is available, they'll bring you from the waiting room to a breakout room (based on the order in which you joined the waiting room).

The 30-minute Guideline

If you get stuck on a homework problem, come by office hours or post on Piazza! We recommend you spend about 30 minutes trying to figure out a problem, and then ask for help. Enough time that you can try a few things to get unstuck, but not SO much time that you're banging your head against the wall. Try for 30 minutes, then ask us. :)

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	7	75%

HW8 is a second-chance		
Quizzes (lowest dropped)	8	20%
Mini-Visualization	1	5%
	TOTAL	**100%**

All quizzes and homeworks have the same weight regardless of the number of points allocated. The percentage score of each is computed in your final average -- for example, if you get 9/11 on Quiz 1, and 9/10 on Quiz 2, then the average of those two is (9/11 + 9/10) / 2 = (81.8% + 90%) / 2 = 85.9%.

You'll submit homeworks and quizzes through Gradescope.

Letter Grades

Your final grade for DS2000 will use the following breakpoints to convert from letter to number grades. We use natural rounding to get these whole numbers, e.g., 96.5 becomes a 97 but 96.4 becomes 96.

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 Rubric & Styleguide. We'll also share grading notes specific to each individual homework. Make sure you keep the rubric and those notes handy as

you're working on an assignment.

The final assignment of the semester, Homework 8, will be a second-chance 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, available Tuesday 4:30pm through Friday 9:50am. The quizzes are auto-graded and gradescope will confirm when you get an answer correct. You can resubmit your answers right up until the quiz deadline.

Mini-Visualization

Data scientists are great communicators. This assignment is a chance to create a visualization and its explanation, and get feedback on your work that has nothing to do with your code.

You'll generate a plot based on data that we assign. You can choose what type of plot, and how you want to share the data. You'll write a short paragraph about your plot and the data, and you'll be graded on (1) the clarity and quality of the plot, and (2) the clarity and quality of your writeup. We will not look at your code.

Late/Makeup Policy

- ***Homework*** - You can submit homeworks up to 24 hours late with no penalty, and 24-48 hours late with a 5% 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-Visualization*** - You must submit this assignment on time. We will not accept any late submissions.

Software

We'll be using Python 3.9 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" DS2000 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) is the best tool for specific questions or concerns about your experience in class or anything sensitive in nature. During the week, I'll respond within 24 hours, but don't expect a response after 9pm. On the weekends I'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.

You can also ask lecture-related questions directly to Laney so I can wrap them into the next lecture if appropriate. Please use this form to do so: https://forms.gle/8i5vkS5dnVKi3PiF9.

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.