

DS2000 -- Programming with Data
Khoury College, Northeastern University
Spring 2020
Dr. Laney Strange (laney@northeastern.edu)

Course web page	http://course.ccs.neu.edu/ds2000
Piazza	https://piazza.com/northeastern/spring2020/ds2000
Lecture Schedule	Sec 1 TF 9:50-10:55am. Internat'l Village 019. Sec 2 TF 1:35-2:40pm. Hurtig 129
Computer Science Practicum	CS Practicum. R 9:50-11:30am WVH 210.
Science Practicum	Science Practicum 1. W 11:45am-1:25pm. Ryder Hall 154 Science Practicum 2. R 11:45am-1:25pm. WVH 210.
Social Science Practicum	SS Practicum 1. W 9:50-11:30am. WVH 210. SS Practicum 2. W 2:50-4:30pm. Hastings 103.
Health Practicum	Health Practicum. R 2:50-4:30pm. Ryder Hall 268
Business Practicum	Business Practicum 1. W 11:45am-1:25pm. WVH 210 Business Practicum 2. W 2:50-4:30pm. Hastings 206 Business Practicum 3. R 11:45am-1:25pm. Ell Hall 411 Business Practicum 4. R 2:50-4:30pm. Ryder Hall 269

Every student registered for DS2000 attends lecture. You must also sign up for a practicum (DS2001).

Instructors

- Laney Strange (lecture). laney@northeastern.edu
- John Rachlin (CS Practicum). j.rachlin@northeastern.edu
- Taylor Braswell (SS Practicum). braswell.t@husky.neu.edu
- Farzaneh Nekui (SS Practicum). nekui.f@husky.neu.edu
- Shun-Yang Lee (Business Practicum). sh.lee@northeastern.edu
- Vance Blankers (Science Practicum). v.blankers@northeastern.edu
- Brecia Douglas (Health Practicum). br.douglas@northeastern.edu

Office Hours (Instructors)

- Laney Strange. WVH 310. TBD

Teaching Assistants

- Allyson Young young.all@husky.neu.edu
- Amrita Suresh suresh.am@husky.neu.edu

- Eric Chung chung.er@husky.neu.edu
- Gaurid Dandi, dandi.g@husky.neu.edu
- Harish Ramani ramani.h@husky.neu.edu
- Heer Patel patel.hee@husky.neu.edu
- Jackson Steilberg steilberg.j@husky.neu.edu
- Kevin Liang liang.ke@husky.neu.edu
- Kristi Bui bui.k@husky.neu.edu
- Parth Barhanpurkar barhanpurkar.p@husky.neu.edu
- Sapna Sharma sharma.sap@husky.neu.edu
- Smit Kiri kiri.s@husky.neu.edu
- Somya Bhargava bhargava.so@husky.neu.edu
- William Cunningham cunningham.wil@husky.neu.edu
- Zhiyuan Cao cao.zhi@husky.neu.edu

Below is the **tentative** office hours schedule for TAs, which is subject to change. We'll keep Piazza and the course website updated. Office hours begin January 10th.

Mon	Tue	Wed	Thu	Fri	Sat	Sun
	5-8pm (RY 299)	1-4pm (RY 141) 4:30-7:30pm (RY 147)	5-8pm (RY205)			1-4pm (RY 141)

Required Textbooks

- Think Python: How to Think Like a Computer Scientist. Allen B. Downey. O'Reilly Media, 2015. ISBN: 1491939362. Available for [download](#) or [purchase](#).
- Matplotlib Reference. Available for [download](#).

Course Description

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 and Excel. 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. Introduces predictive analytics for forecasting and classification. Demonstrates the limitations of statistical techniques.

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)	Topics
Downey Ch. 2	Variables, data types, arithmetic operations
Downey Ch. 3, 6	Functions
Downey Ch. 5	Conditionals, boolean expressions
Downey Ch. 7	Iteration (while loops, for loops)
Downey Ch. 8	Strings
Downey Ch. 10	Lists
Downey Ch 11	Dictionaries and JSON
Downey Ch 15	File Processing -- unstructured data, CSV files
	Application Programming Interfaces (APIs)
Matplotlib ch 1-4	Visualization with Matplotlib

Evaluation

Factor	Number	Weight
Homework Sets	6	30%
Practicum	weekly	20%
Midterm Exam	1	15%
Quizzes	4	10%
Final Project + Presentation	1	25%
	TOTAL	100%

You'll submit homeworks through the Khoury HandIn server. Allow at least one week after you submit a homework or exam before the grades are posted. If you have a question about a grade or would like a score to be reviewed, please come by office hours so we can discuss in person.

Quizzes / Exams

Four quizzes will be given this semester. They will be administered during the first 15 minutes of class. You must be present to receive a grade for each quiz.

There are 5-7 questions per quiz. Your quiz grade will be scaled, though (for example, getting one question wrong on a 6-question quiz doesn't mean your quiz score is $5/6 = 83\%$). Quiz scaling will be applied as follows:

- Zero incorrect: Perfect
- One incorrect: Good
- Two incorrect: Satisfactory
- Three incorrect: Fair
- Three or four incorrect: Unsatisfactory
- More than four incorrect: Poor

There is one midterm exam, given about halfway through the semester. It will be administered during the lecture period. Exam score will be out of 100 points.

Homework Sets

Homeworks are assigned (almost) every week until we move on to the project. They are due approximately one week after they are assigned, unless otherwise noted.

Homeworks will be evaluated according to the [DS2000 Grading Rubric](#).

You are allocated two "late date" passes; each of which gives you an additional 24 hours to submit a homework without penalty. You can apply both passes to a single homework or split it between two homeworks. A late day pass grants you 24 hours and cannot be broken into smaller chunks.

Project

The goal of the project is to gain hands-on experience with finding, importing, analyzing, visualizing, and presenting a dataset of your choosing. You can work alone or in a small group (of 2--4 members) -- you will first submit this group's topic, membership, and division of labor in a proposal.

At the conclusion of the class you will submit your Python code along with any datasets you used in the project. Additionally, you and your group will present your work during the last 1.5 weeks of the semester, either in lecture or in practicum.

Technical Requirements

We'll be using Python 3 in this class. We'll use IDLE, Python's own Integrated Development Environment (IDE). An IDE combines the Python interpreter with an editor for your code, which makes it easy to work on your code and test/run your software.

Install IDLE on your own laptop and make sure you bring it to practicum. Download Python 3.8 from <https://www.python.org/downloads>. It's available for Windows, Mac OSX, and Linux. Once installed, click on IDLE to open it up. You can use Python's interactive environment, or you can write and save a file with a .py extension.

You also must sign up for a Khoury account, so that you have access to the system we use for homework

submissions and grading. Follow these instructions to register for your Khoury account: bit.ly/ccisaccount

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.

Late/Makeup Policy

All assignments are expected to be completed and turned in on schedule. Due dates will be clearly indicated for each assignment.

You must be present in lecture to receive a grade for quizzes and exams. If you must miss a quiz/exam due to extreme, unanticipated circumstances such as an illness or a family emergency, notify me via email before the event.

Attendance Policy

Lecture attendance is not required for lectures, but it is for quizzes, exams, and practicum sessions.

It is your responsibility to familiarize yourself with the course schedule to ensure that you do not have any conflicts with important dates. If you must miss a lab or exam, and if you feel that extraordinary circumstances warrant a makeup, get in touch with me before the scheduled date.

When you come to class, I ask that you be fully present. No phones are permitted in the classroom. If you use a laptop, use it **only to take notes**. Please be respectful of your fellow students and me by participating attentively and non-disruptively.

Academic Integrity

While students are encouraged to discuss course materials, no plagiarism/copying is allowed on

homework. In particular,

- You may not copy anyone else's code under any circumstances.
- You may not permit any other student to see any part of your program, except when requesting assistance in debugging.
- You may not permit yourself to see any part of another student's program, except when rendering assistance in debugging.
- You may not post a public question to Piazza that contains any part of your code.

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](#).