DS2000 -- Programming with Data
Northeastern University CCIS
Spring 2019

Course web page	http://course.ccs.neu.edu/ds2000
Piazza	https://piazza.com/northeastern/spring2019/ds2000
Lecture Schedule	Sec 1 TF 9:50-10:55am. BK010. Sec 2 TF 1:35-2:40pm. RI 236.
Computer Science Practicum	CS Practicum 1. W 11:45am-1:25pm. WVH 212. CS Practicum 2. W 2:50-4:30pm. WVH 210.
Science Practicum	Science Practicum 1. W 2:50-4:30pm. WVH 212. Science Practicum 2. W 4:40-6:20pm. WVH 212.
Social Science Practicum	Social Science Practicum. W 11:45am-1:25pm. WVH 210.
Health Practicum	Health Practicum. T 3:25-5:05pm. RY 128.

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

Instructors

- Laney Strange (lecture). laneys@northeastern.edu.
- Stephen Intille (practicum). s.intille@northeastern.edu.
- Samuel Judge (practicum). s.judge@northeastern.edu.
- Piotr Sapieżyński (practicum). sapiezynski@gmail.com.
- Sarah Shugars (practicum). shugars.s@husky.neu.edu.

Office Hours (Instructors)

- Laney Strange. Tuesday, 3:00-6:00pm. WVH 310.

Office Hours (TAs)

Teaching Assistants for this course are:

- Will Slotterback. slotterback.w@husky.neu.edu
- Edward Wersocki. wersocki.e@husky.neu.edu
- Jackson Steilberg. steilberg.j@husky.neu.edu
- Anna Rojek. rojek.a@husky.neu.edu.
- Jake Mittleman. mittleman.j@husky.neu.edu
- Darshana Jaint. jaint.d@husky.neu.edu
- Magnus Frenberg. frennberg.m@husky.neu.edu
- Mahita Valluru. valluru.m@husky.neu.edu

Office hours are scheduled throughout the week. We're still finalizing the schedule as we go to print with the syllabus, and we'll keep Piazza and the course website updated.

Mon	**Tue**	**Wed**	**Thu**	**Fri**	**Sat**	**Sun**
3-6pm (KA204)		9am-12pm (SL047) 12-3pm (KA204) 6-9pm (KA208)	1-4pm (SL009) 6-9pm (KA202)			

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.

No prior programming experience is assumed; therefore, suitable for students with little or no computer science background.

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
Downey Ch 12	Tuples

Matplotlib ch 1-4	Visualization with Matplotlib
	Structured Data (CSV; JSON)
	Application Programming Interfaces (APIs)

Evaluation

Factor	**Number**	**Weight**
Homework Sets (lowest dropped)	5	30%
Practicum	weekly	20%
Midterm Exam	1	15%
Quizzes	4	10%
Final Project + Presentation	1	25%
	TOTAL:	**100%**

Your lowest-scoring homework assignment will be discarded and not counted toward your final grade.

You can check your homework grades online through the CCIS 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 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 may redeem one (only one!) homework late during the semester. This is your "late token." If you wish to cash in it, email laneys@northeastern.edu no later than one hour before the original deadline. You will then have an additional 3 days to complete and submit the assignment.

Apart from your late token (and your one dropped homework), no late homeworks will be accepted. You will receive zero credit for assignments submitted after the deadline.

Homework Sets will be posted on the course website. You will submit your homework solution via the CCIS HandIn server; instructions for submission will be included in the homework documentation.

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--3 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 will present your work during the last week of class.

Technical Requirements
We'll be using Python 3 in this class.

The rooms for our scheduled practicum sessions have desktops with Python 3 installed on them. You should also download Python 3 onto your own computer before the semester begins. 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.

Download Python 3.7.1 from https://www.python.org/downloads/release/python-371/. 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 CCIS account. Follow these instructions to register for one: bit.ly/ccisaccount

Communication
Computer Science is equal parts art and science. There is rarely a problem to solve for which only one

solution exists. Computer scientists develop good software by applying knowledge, educated guesses, trial-and-error, and collaboration. We have office hours for DS2000, but it is often just as helpful to talk over your approach with your classmates as it is to talk it over with a Teaching Assistant or Professor.

The quickest way to get feedback and help 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, cashing in your late token, 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 computer 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. Apart from your "late token," homeworks submitted after the deadline will not be accepted.

There is no late token for quizzes or exams. You must be present to receive a grade. 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.