DS 2000: Programming with Data, Fall 2022

Instructors	Felix Muzny[1] (pronunciation: "Muse-knee"; pronouns: they/them and he/him) Contact: f.muzny@northeastern.edu John Rachlin[2] (pronunciation: "RACK-lin; pronouns: they/them and he/him) Contact: j.rachlin@northeastern.edu
Course web page	http://course.ccs.neu.edu/ds2000
Office Hours	See course website
Piazza	https://piazza.com/class/l5id83u7mlg4fg
Gradescope	https://www.gradescope.com/courses/408909 Gradescope pulls its roster from Canvas, but not automatically, so be patient if you've just registered for DS 2000 and aren't in Gradescope yet.
Lecture Schedule	Sec 1: TF 11:45am - 12:50pm. Live/Online only. (Lectures are not recorded.) Rachlin. Sec 2: TF 9:50am - 10:55am. RI 236. Muzny. Sec 3: TF 1:35pm - 2:40pm. RI 236. Muzny. Sec 4: TF 3:25pm - 4:30pm. CH 103. Muzny.
Practicum Schedule (DS2001)	1. R 11:45am - 1:25pm. Online only. Rachlin. Data Science. 2. W 9:50am - 11:30am. WVH 212. Sathyanarayana. Health Science. 3. W 9:50am - 11:30am. WVH 210B. Mosca. Data Science. 4. W 9:50am - 11:30am. KA 005. Zhang. Social Science & Humanities. 5. W 11:45am - 1:25pm. WVH 210A. Zhang. Social Science & Humanities. 6. W 11:45am - 1:25pm. WVH 212. Luo. Business Admin. 7. (no section 7) 8. W 2:50 - 4:30 pm. WVH 210A. Gillani. Business Admin. 9. W 2:50 - 4:30 pm. WVH 212. Luo. Business Admin. 10. W 2:50 - 4:30 pm. WVH 210B. Uslu. Social Science & Humanities. 11. W 4:40 - 6:20 pm. WVH 210A. Gillani. Business Admin. 12. (no section 12) 13. R 9:50am - 11:30am. WVH 210B. Strange. Data Science. 14. R 9:50am - 11:30am. WVH 212. Mosca. Data Science. 15. R 11:45am - 1:25pm. WVH 210A. Strange. Data Science. 16. R 11:45am - 1:25pm. WVH 212. Mosca. Data Science. 17. R 11:45am - 1:25pm. WVH 210B. Luo. Business Admin. 18. R 2:50 - 4:30 pm. WVH 212. Qu. Social Science & Humanities. 19. (no section 19) 20. R 2:50 - 4:30 pm. WVH 210B. Luo. Business Admin.
TAs	See course website TA page for TAs & their office hours!

[1] Call them "Felix", "Professor Muzny", or "Professor Felix"
[2] Call them "Professor Rachlin"

Fall 2022 Course Format

DS2000/2001 are organized as in-person, on-ground classes for Fall 2022 unless you are specifically registered for Section 1 of DS 2000 and/or Section 1 of DS 2001. This means that we expect you to attend class in person if you are well. We expect all members of the DS 2000 community to individually monitor themselves for illness. Do not attend DS 2000 lecture in person if you are ill in any way.

Professor Felix is immunocompromised and **will be masking** for the duration of the semester during lecture. (And would prefer that you do the same).

If you are in Section 2, 3, or 4 and unable to attend class on any given day, your first, best option is to attend Prof. Rachlin's lecture (T/F from 11:45am - 12:50pm - See canvas for Zoom link information). We will not be recording lectures but we will be posting all materials from class after lecture. We encourage you to carefully walk through these materials and come to office hours to make sure that you don't fall behind. Please reach out if you feel you are falling behind—we have a large team to help!

In general, if you can't attend the lecture for the section that you are enrolled in on a specific date (this should not be a regular occurrence), you are welcome to attend a different section on that day.

All members of the DS 2000 community will follow university guidelines regarding masking, social distancing, and other public-health related policies.

DS 2001: Consult your individual DS 2001 guidelines for protocols about what to do if you are unable to attend DS 2001 on any given week.

Course Description - DS 2000

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.

New 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

Course Description - DS2001

You may wonder "what's going on with DS 2001?". DS 2001 is your lab that runs in parallel to this course. The different sections of DS 2001 are different and may focus on different subject material. You can expect your DS 2001 course instructor to provide all details of the work and projects that you'll complete for DS 2001 and how you'll be evaluated in that course. You **will receive separate grades** for DS 2000 and DS 2001.

TA Office Hours

We have office hours all week long. Please check the DS2000 website for the schedule, and keep an eye on Piazza for any updates. Remember that you are welcome to attend the office hours of any TA, not just the ones assigned to your section of DS 2001!

TAs are here to help and happy to answer your questions about the course material, homework assignments, practicum, and the project. We expect you to treat TAs with kindness and respect; we expect the same of your behavior towards one another and towards all the instructors.

Come to office hours well before an assignment is due; they often get crowded, and if you arrive a few hours before an assignment is due, you'll end up feeling rushed and frustrated. Even if you find yourself in that unhappy situation, bear in mind that it is not the TAs' job to provide the solution for you. We host office hours so we can debug **your understanding of the material**, not your code.

Recommended Textbooks

Downey (2016). *Think Python 2nd Ed.* Available as a free PDF, HTML, or Print.
This book is a terrific guide to learning to program in Python, particular if you are new to programming.

Deitel and Deitel (2019). *Intro to Python.* Available for free online through O'Reilly/Safari or in Print.
This book is more of a reference. It is quite comprehensive and uses the Anaconda Python distribution as we do in this course.

Felix's Python Notes available online (https://muzny.github.io/csci1200-notes/intro)

Software and Communication

In this class, we'll be using Anaconda Python distribution based on Python 3.x. Before our first meeting of DS 2000, you'll work with your DS 2001 instructor to download and set up everything you need to program with Python for your practicum assignments, homeworks, and the project.

The main piece of software we'll be using is:
- Anaconda Navigator

- Spyder (included with Anaconda Navigator installations)

You will be submitting your assignments on gradescope.

There is one Piazza page for all of DS2000/2001: https://piazza.com/class/l5id83u7mlg4fg

You can always reach out to course staff via email or in office hours.

Classroom Environment & Expectations

- **Attendance:** You are strongly encouraged to attend lecture synchronously whenever possible. If you are unable to attend lecture on a given day, it is your responsibility to attend office hours, consult lecture materials, and follow-up as needed.
- **Classroom environment:** It is unusually common in Computer Science or Data Science classes for some students to ask questions that are not really questions so much as opportunities to demonstrate knowledge of vocabulary or facts beyond the topic at hand. This can have a discouraging effect on other students who are not familiar with those terms, causing them to worry that they are less prepared to do well in the class (this is rarely the case—knowing terms outside the scope of the course is not a good predictor of success). If you find yourself wanting to make such a question or comment, please come talk to your instructor about the topic after class or during office hours—we're always happy to discuss tangentially related topics at those times!
- **Accommodation letters:** If you have an accommodation letter, please bring it to your instructor at your earliest convenience so that we can make sure this class is meeting your needs.
- **Name and pronouns:** If your name and pronouns are not in alignment with those listed on our class roster, please let us know either in person or via email so that we can ensure you are correctly addressed in this class.
    - If you wish to add, change, or update your pronouns in Canvas, see this knowledge base webpage. To update your display name in Canvas, see this article.
- **Class expenses:** If obtaining any material for use in our class presents a financial hardship for you, please let me know and I will work with you to locate the resources that you need to succeed in this class.
- **Feedback:** Please don't hesitate to reach out to me if any aspect of this course or class community could be improved.

Schedule

Lectures run on Tuesday and Fridays. Please make sure to attend the lecture that you've registered for!

Below is the tentative schedule. More details will be released by-week on the course website.

Week	Dates	Quiz	HW	Topics
1	Note! You **will** have DS 2001 on Sept 7/8 Sept 9		HW 1 (due 9/16)	Introduction, variables, assignment, arithmetic operations
2	Sept 13, 16	Quiz 1	HW 2 (due 9/23)	intro to files, reading text and numbers, data viz
3	Sept 20, 23	Quiz 2	HW 3 (due 9/30)	boolean expressions, conditionals, iteration (while

				loops), math & random libraries
4	Sept 27, 30	Quiz 3	HW 4 (due 10/7)	data sources, lists, iteration (for loops)
5	Oct 4, 7	Quiz 4	HW 5 (due 10/14)	lists, working with multiple lists, functions
6	Oct 11, 14	Quiz 5	HW 6 (due 10/21)	functions, parameters and scope, returns, csv files, 2d lists
7	Oct 18, 21	Quiz 6	HW 7 (due 10/28)	2d lists, data viz, tuples
8	Oct 25, 28		HW 8 (due 11/4)	dictionaries, counting occurrences
9	Nov 1, 4	Quiz 7	HW 9 (due 11/18)	string processing, sentiment analysis, cleaning data
10	Nov 8, no class on Nov 11 (Veteran's Day)			intro to classes and objects Veteran's day is Nov 11th this year. We **do** have lecture on Nov 8th.
11	Nov 15, 18	Quiz 8		classes and objects
12	Nov 22, no class on Nov 25 (Thanksgiving)			JSON formats, APIs, requests library Thanksgiving recess is Nov 23 - 27th this year. We **do** have lecture on Nov 22nd.
13	Nov 29, Dec 2	Quiz 9		Pandas, data frames
14	Dec 6			wrap up

Evaluation

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

Factor	Number	Weight
Homework Sets (**none** dropped)	9	90%
Quizzes (**lowest one quiz** dropped)	9	10%

Final grades for DS 2000 will be based on the following scale. We will round all grades to the nearest whole number (e.g. 94.5 will become 95, 94.4 will become a 94):

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 are due on **Fridays at 9pm**. Please contact your DS2001 practicum instructor for practicum assignment due dates.

Homeworks will be evaluated according to the DS2000 Grading Rubric, posted on the course website.

Quizzes

We'll have weekly quizzes, released on Tuesdays at 4:30pm. 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 at 9:50am on Fridays.

Your lowest quiz grade will be dropped.

Late Policy

Life happens. You may turn in any homework assignment (but not quizzes) up to 2 days late (until Sundays at 9pm) for a 10% penalty. If you were to receive a 95% on a homework but it is submitted late, you would receive an 85% instead. You do not need to do anything special to request this late submission.

In the case of extenuating circumstances, please reach out directly to course staff as soon as possible to discuss your individual circumstances. Proactive communication is always preferred whenever possible.

Academic Integrity

Homework for DS 2000 is expected to be completed as individuals. This means that you should not receive code in any form (written down, sent pictures, be told aloud) from other people. If in doubt, we urge you to come to office hours. Code is both extremely individual (which can be surprising to beginner programmers) and It is extremely hard to "unsee" once you have seen a solution, even if you are trying to look just for reference.

Violations of the academic integrity policy will result in a 0 on the given assignment for all parties involved. Repeated violations may result in a report to the Office of Student Conduct and Conflict Resolution (OSCCR.)

The university's academic integrity policy discusses actions regarded as violations and consequences for students: http://www.northeastern.edu/osccr/academic-integrity

Accommodations

It is **our job** to create a classroom environment that is most conducive to you learning well. If you have accommodations from the [Disability Resource Center](#), please provide your letter to me early in the semester so that I can arrange for these accommodations. If you wish to receive accommodations and do not have a letter, please visit the DRC at 20 Dodge Hall or call (617) 373-2675.

Classroom Environment

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. You are encouraged to disagree with other students and the instructor, but such disagreements need to respectful and be based upon facts and documentation (rather than prejudices and personalities). The instructor reserves the right to interrupt conversations that deviate from these expectations. Repeated unprofessional or disrespectful conduct may result in a lower grade or more severe consequences.

Part of the learning process in this course is respectful engagement of ideas with others.

The [Code of Student Conduct can be found on the OSCCR website](#).

Title IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking.

The Title IX Policy applies to the entire community, including male, female, non-binary, and transgender students, faculty and staff.

If you or someone you know has been a survivor of a Prohibited Offense, confidential support and guidance can be found through [University Health and Counseling Services](#) staff and the [Center for Spiritual Dialogue and Service](#) clergy members.

By law, those employees are not required to report allegations of sex or gender-based discrimination to the University.

Reports can be made non-confidentially to the Title IX Coordinator within the Office for Gender Equity and Compliance at: titleix@northeastern.edu and/or through NUPD (Emergency 617.373.3333; Non-Emergency 617.373.2121).

Reporting Prohibited Offenses to NUPD does NOT commit the victim/affected party to future legal action.

Faculty members are considered "responsible employees" at Northeastern University, meaning they are

required to report all allegations of sex or gender-based discrimination to the Title IX Coordinator.

In case of an emergency, please call 911.

Please visit http://www.northeastern.edu/titleix for a complete list of reporting options and resources both on- and off-campus.

Religious Holidays

The course staff will make every effort to deal reasonably and fairly with all students who, because of religious obligations, have conflicts with scheduled exams, assignments or required attendance. In this class, contact the course staff at least 7 days in advance of the conflicting date to reschedule a practicum or project due date.