CHEM 4803DR/8843DR Syllabus — Fall 2019

Special Topics: Python for Data Science

Lecture: TR 12–1:15, Architecture West 259

Instructor and TA Information

	E-mail	Office Location	Office Hours
Instructors			
Prof. Robert Dickson	dickson@chemistry.gatech.edu	MoSE G209A	R 2:00-3:00p
Mr. Dominic Sirianni	sirianni.dom@gatech.edu	MoSE 4202A	W 2:00-3:30p
Teaching Assistant			
Baijie Peng	baijiepeng@gatech.edu	MoSE G209A	TBA

1.1 General Information

Course Description

During the past several years, Data Science has become one of the most rapidly growing fields in all of computing. From advertising targeting individual users on Facebook to predicting future stock market behavior, the principles of Data Science can be applied to solve problems in an incredible diversity of contexts. In this course, students will learn the basic principles of Data Science and develop skills working with the most common tools in the world of Data Science, building from foundational experience with computer programming in the highly versatile Python language. The knowledge and skills developed in this course will therefore be transferable directly to students' future careers in the science, technology, or business sectors.

Pre- and/or Co-Requisites

This course is intended for upper-level undergraduate and graduate students in Chemistry & Biochemistry, who want to develop their knowledge and skills in Data Science and Python programming. No background in computer programming in Python or any other language is required — the only requirement is the willingness to learn!

Other useful background knowledge: Calculus; Introductory Linear Algebra

Course Goals and Learning Outcomes

Upon successful completion of this course, students will be able to...
  (i) Employ software development best practices to implement efficient, readable Python programs,
 (ii) Apply machine learning/data science principles to solve real-world problems using relevant tools,
(iii) Utilize various Python libraries common in data science applications, and
 (iv) Employ version control and social coding practices to work effectively with peers on collaborative software projects.

1.2 Course Requirements and Grading

The most effective method to learn to apply new conceptual knowledge using new tools to solve problems is to spend time actually solving problems. Therefore, there will be no exams or quizzes in this class; rather, this course will be entirely project-based. Typically, project assignments will consist of:

(i) a mini-project (to be completed during class), which provides a primer for students to become acquainted with the use of a particular concept or tool, and

(ii) the project proper (to be completed outside class), where students will write Python software to solve a real-world or relevant problem using the concept or tool.

Projects will be assigned throughout the semester, with approximately two weeks between the completion of the mini-project in class and the due date for the project itself. A capstone project which combines knowledge and skills acquired throughout the semester, together with subsequent demo/presentation, will be completed by each student in lieu of a final exam. Project topics will be chosen by the student either from their own research questions or from a list provided by the instructors.

Assignment Submission & Code Review

Assignments will be turned in by each individual student via submission of a pull request (PR) to the class GitHub repository (see "Course Website" section below for more details). Upon submission, each assignment will recieve a letter grade according to the criteria below:

A: High/Very high quality work
B: Good quality work
C: Substandard/poor quality work

In addition to a letter grade, each assignment submission will be subject to interactive PR "code review" from the instructor(s)/TA on GitHub (see "Course Website" section below) where any potential room for improvement will be identified and relevant suggestions made. Through active participation in the code review process (in particular, by improving their assignment submissions according to the suggestions made in code review), students can improve the letter grade of *any* assignment submission up to A-quality. Participation in code review and any improvements to assignment submissions must be completed within one week of the original assignment's due date, at which time the PR will be closed and no further changes will be accepted.

Grading Scale

Your final grade will be assigned a letter according to a weighted average of the quality of work submitted throughout the course. Special consideration will be given, however, to improvement throughout the semester — therefore, as long as each student continues to improve during the course, there's no reason why everyone shouldn't earn an A!

Extra Credit and Grade Dispute Policies & Procedures

Due to the opportunity to improve the letter grade for any assignment submission via participation in code review, no extra credit for individual assignments will be given. Additional extra credit may be made available at the sole descretion of the instructors.

1.3 Course Materials

Course Text

1. J. VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data* (O'Reilly, Boston, MA, 2017).

Note: This book should be accessible via the web at `https://learning.oreilly.com/library/view/python-data-science/9781491912126/` for free with your GaTech student credentials.

Course Website and Other Classroom Management Tools

We will use Canvas as our primary course website. This can be accessed by signing into your account at `www.canvas.gatech.edu`. Occasionally throughout the semester, additional notes, videos, solutions, or other relevant resources will be posted to the Canvas website. You will be notified in advance whenever these materials are to be made available and will therefore be responsible for their content, unless otherwise indicated.

For assignment submission and code review, we will utilize a course GitHub repository at `https://github.gatech.edu/dsirianni6/Python-Data-Science`. Materials to be utilized in class will pushed to this repository in advance of the class meeting, and an announcement will be made on Canvas to notify students to sync their forks of the repository.

1.4 Course Expectations and Guidelines

Academic Integrity

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. For information on Georgia Tech's Academic Honor Code, please visit `http://www.catalog.gatech.edu/policies/honor-code/` or `http://www.catalog.gatech.edu/rules/18/`. Any student suspected of cheating or plagiarizing on a quiz, exam, or assignment may be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations.

Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404) 894-2563 or `http://disabilityservices.gatech.edu/`, as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter. Please also e-mail the instructors as soon as possible in order to set up a time to discuss your learning needs.

Attendance and Participation

Other than when required by the Institute, *attendance will not be taken.* Instead, classroom participation will be determined through "formative assessments" (FAs) during class. While these FAs are not graded for correctness, you should use them as a tool to assess your own learning by (i) identifying areas or concepts with which you are struggling and (ii) to gauge your progress and improvement as you study those concepts in particular. Just as you will be expected to use these assessments to guide your learning, the instructors will be using the collective results from these assessments to guide lesson planning and our classroom approach. This way, we can each work to be the best students and instructors we can be, without waiting until the end of the semester to analyze the results of our efforts.

The instructors do understand, however, that life happens — your car may break down, you may get sick, or some other event may occur which prevents you from attending class. Despite the fact that attendance is not taken nor is classroom participation graded, however, we (the instructors) strongly suggest only to miss class for excused absences, defined by the Institute as "documented reasons of illness, family emergency, or participation in approved Institute activities" (please see `http://catalog.gatech.edu/rules/12/` for more information). We make this recommendation because we design each classroom activity, discussion, and experiment to address a specific goal: to help you make measurable progress towards our course learning objectives. Therefore, your attendance at *all* class meetings (including lecture and any recitation or review session(s)) and active participation in these activities throughout the semester will be critical for your success in this course.

Collaboration and Group Work

Effective education is collaborative by nature. We encourage you to consult your fellow students, TA, instructors, other faculty, and/or staff throughout this course. However, all assignment submissions must be your own! This means that you must create your own derivations, solutions, analysis, thoughts, justifications, data tables, results, plots, calculations, etc. on your own. Copying and pasting figures, tables, or text, even from group partners, constitutes plagiarism and is a violation of the GT Honor Code.

The TA is bound by the Honor Code to notify the instructors of any suspected plagiarism of any assignment. If you are concerned about potential plagiarism associated with one of your assignment submissions, please consult your TA — do this before submitting your assignment! The instructors will initially address all Honor Code violations; a violation will only be referred to the Office of Student Integrity if a mutual agreement is not reached. For more information, please see the Georgia Tech Academic Honor Code (`http://osi.gatech.edu/plugins/content/index.php?id=46`).

Extensions, Late Assignments, and Re-Scheduled/Missed Exams

Unless expressly indicated otherwise, all assignments may be submitted on GitHub up to the start of class on the due date. Assignments submitted more than six (6) calendar days late will not be eligible for grade improvement by code review, and any assignment submitted more than seven (7) calendar days late will not be accepted. Regardless of the timeliness of assignment submission, the window for participation in code review will only extend to one week (7 calendar days) following the assignment's original due date. Therefore, for every day an assignment is late, one fewer day is available for code review and assignment

Extensions on assignments will be given only for documented reasons of illness, family emergency, or participation in approved Institute activities (such as field trips and athletic events, see `http://catalog.gatech.edu/rules/12/` for more information). Presentations and/or demonstrations of capstone projects, scheduled to be given during the final exam period for the course, will be re-scheduled in accordance with Georgia Tech's policy, also found at `http://www.catalog.gatech.edu/rules/12/`.

Student-Faculty Expectations Agreement

At Georgia Tech, we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See `http://www.catalog.gatech.edu/rules/22/` for an articulation of some basic expectations that you can have of me and that I have of you. In the end, simple respect for knowledge, hard work, and cor-

dial interactions will help build the environment we seek. Therefore, you are encouraged to remain committed to the ideals of Georgia Tech while in this class.

Student Use of Electronic Devices in the Classroom

Electronic devices (e.g., tablets/laptops) are *required* for each classroom meeting, as they are necessary for classroom participation. However, use of these electronic devices for other purposes during class (e.g., Facebook, texting, phone calls, etc.) disrupt the classroom learning environment for all other students, and therefore will not be tolerated.

1.5   Campus Resources for Students

In your time at Georgia Tech, you may find yourself in need of support. Below you will find some resources to support you both as a student and as a person.

I. Academic support:
- Center for Academic Success (`http://success.gatech.edu`)
  - 1-to-1 tutoring (`http://success.gatech.edu/1-1-tutoring`)
  - Peer-Led Undergraduate Study (PLUS) (`http://success.gatech.edu/tutoring/plus`)
  - Academic coaching (`http://success.gatech.edu/coaching`)
- Residence Life's Learning Assistance Program (`https://housing.gatech.edu/learning-assistance-program`)
  - Drop-in tutoring for many 1000 level courses
- OMED: Educational Services (`http://omed.gatech.edu/programs/academic-support`)
  - Group study sessions and tutoring programs
- Communication Center (`http://www.communicationcenter.gatech.edu`)
  - Individualized help with writing and multimedia projects
- Academic advisors for your major (`http://advising.gatech.edu/`)

II. Personal Support:
- The Office of the Dean of Students: `http://studentlife.gatech.edu/content/services`; 404-894-6367; Smithgall Student Services Building 2nd floor
  - You also may request assistance at `https://gatech-advocate.symplicity.com/care_report/index.php/pid383662?`
- Counseling Center: `http://counseling.gatech.edu`; 404-894-2575; Smithgall Student Services Building 2nd floor
  - Services include short-term individual counseling, group counseling, couples counseling, testing and assessment, referral services, and crisis intervention. Their website also includes links to state and national resources.
  - Students in crisis may walk in during business hours (8am-5pm, Monday through Friday) or contact the counselor on call after hours at 404-894-2204.
- Students' Temporary Assistance and Resources (STAR): `http://studentlife.gatech.edu/content/need-help`
  - Can assist with interview clothing, food, and housing needs.
- Stamps Health Services: `https://health.gatech.edu`; 404-894-1420

1.6 Statement of Intent for Inclusivity

As members of the Georgia Tech community, the instructors are committed to creating a learning environment in which all of our students feel safe and included. Because all of us are individuals with varying needs, we are reliant on your feedback to achieve this goal. To that end, we invite you to enter into dialogue with us about the things we can stop, start, and continue doing to make our classroom an environment in which every student feels valued and can engage actively in our learning community.

1.7 Course Schedule

See rough outline of dates, topics to be covered, and notes used in additional "Course Outline" document on the Syllabus page of Canvas.

1.8 Practical Advice for Success

Programming and Data Science are challenging subjects, but they are extremely rewarding to study thanks to their growing application and influence. Because we am committed to helping you to succeed in this course, we have determined the following general advice (in no particular order) to help you to do so:

Come to class. The instructors for this course put a significant amount of time designing activities, lectures, materials, and laboratories which should help you succeed in this course. The best way for you to take advantage of this is, quite simply, to come to class and participate.

Be curious. Several of the concepts we will be discussing this semester may seem strange and unintuitive. When you get confused, try to remain curious and open to these new concepts — this attitude will help you to persevere and make the learning easier, not to mention more fun!

Be bold. Some problems we will encounter this semester will seem daunting. Whether this is because of the mathematics involved, the complexities of using the various new tools you are being exposed to, or just the challenge of thinking in a completely new way, sometimes all that is required is the courage to attempt to solve these problems using completely new concepts and approaches.

Ask for help. Every person who has ever studied programming or been an active programmer has gotten stuck at some point. For our purposes, that means that at some point in the semester, everyone will need to ask someone else (or StackOverflow) for help. That may mean coming to see me or one of the TAs during office hours, collaborating with your peers, or both — just remember, if you're not confused at some point, you're probably not doing something right!

Welcome to CHEM 4803DR/8843DR!
We are so excited that you are here!