DSA 8640 Programming for Data Science
Fall 2024

GENERAL INORMAITON

Course Web Page (Canvas): https://clemson.instructure.com/courses/238255

Instructor: Hongki Kim (Kenn), Ph. D.
Office Hours: Saturday 7:00 pm – 8:00 pm via Zoom / by appointment via Zoom
Zoom link: https://clemson.zoom.us/j/95548335215
Email: hongkik@clemson.edu

Teaching Assistant: TBA
Office: TBA
Office Hours: N/A
Email: TBA

Course Description

This course builds the skills necessary to use Python to develop larger programs and libraries. Students will learn to design, implement and debug Python functions and programs, including object-oriented techniques. The course covers Python data structures, and Python facilities for working with files, strings, regular expressions, DataFrames, APIs, and webpages.

Learning Outcomes

Upon completion of the course, students will be able to:

- Understand data structures (e.g., list, tuple, dictionary), execution controls (e.g., selection, iteration), and libraries / packages in Python
- Access data from files (e.g., text, CSV, JSON) in Python
- Understand namespaces and classes in Python
- Use Pandas DataFrame for data manipulation and visualization
- Perform web scrapping through Python

Required Textbook and Materials

- *Introduction to Computing Using Python* by Ljubomir Perkovoc, 2nd edition.
- Slides and codes will be available on Canvas, or transmitted electronically.
- Laptop with Python 3.12 environment (e.g., Anaconda)

Course Lecture

Every Monday and Wednesday, the instructor will upload a video lecture on Canvas, by following the course schedule (please see ***Fall 2024 Schedule for DSA 8640*** below). Students are expected to watch video lectures on **a weekly basis**. This course is entirely asynchronous. **Successful students will engage with the course material each week and participate on a regular basis**. Students should expect to spend **at least 3 hours per credit hour** on the course.

Office Hour

If you wish to meet with your instructor, you can use the zoom link on the course web page (Canvas). If you are unavailable during the assigned office hour, you must coordinate an appointment date and time via email.

Statement of Instructor/Student Interaction

Students are expected to interact with each other and with the instructor on a weekly basis. Discussions on course topics will mainly take place in Discussion Board on Canvas.

Email Communication

Because of privacy regulations, University faculty and staff may email students only through Clemson email. Therefore, you must use your Clemson email account in this course for all email communications. Check your Clemson account at least three times per week for important messages.

COURSE POLICY

Academic Integrity

From the "Academic Integrity" policy under "Academic Regulations" in the Clemson University's Undergraduate Announcements found on the Registrar's home page:

> As members of the Clemson University community, we have inherited Thomas Green Clemson's vision of this institution as a "high seminary of learning." Fundamental to this vision is a mutual commitment to truthfulness, honor, and responsibility, without which we cannot earn the trust and respect of others. Furthermore, we recognize that academic dishonesty detracts from the value of a Clemson degree. Therefore, we shall not tolerate lying, cheating, or stealing in any form.
>
> When, in the opinion of a faculty member, there is evidence that a student has committed an act of academic dishonesty, the faculty member shall make a formal written charge of academic dishonesty, including a description of the misconduct, to the Associate Dean for Curriculum in the Office of Undergraduate Studies. At the same time, the faculty member may, but is not required to, inform each involved student privately of the nature of the alleged charge.

Consistent with Clemson University's Undergraduate Announcements acts of plagiarism may be resolved through the Plagiarism Resolution Form.

Student Accessibility Services

Clemson University values the diversity of our student body as a strength and a critical component of our dynamic community. Students with disabilities or temporary injuries/conditions may require accommodations due to barriers in the structure of facilities, course design, technology used for curricular purposes, or other campus resources. Students who experience a barrier to full access to a class should let the professor know and make an appointment to meet with a staff member in Student Accessibility Services as soon as possible. You can make an appointment by calling 864-656-6848 or by emailing studentaccess@lists.clemson.edu. Students who receive Academic Access Letters are strongly encouraged to request, obtain and present these to their professors as early in the semester as possible so that accommodations can be made in a timely manner. It is the student's responsibility to follow this process each semester. You can access further information here: http://www.clemson.edu/campus-life/campus-services/sds/.

Title IX (included as required by the University)

The Clemson University Title IX (Sexual Harassment) Statement: Clemson University is committed to a policy of equal opportunity for all persons and does not discriminate on the basis of race, color, religion, sex, sexual orientation, gender, pregnancy, national origin, age, disability, veteran's status, genetic information or protected activity (e.g., opposition to prohibited discrimination or participation in any complaint process, etc.) in employment, educational programs and activities, admissions and financial aid. This includes a prohibition against sexual

harassment and sexual violence as mandated by Title IX of the Education Amendments of 1972. This policy is located at http://www.clemson.edu/campus-life/campus-services/access/title-ix/. Mr. Jerry Knighton is the Clemson University Title IX Coordinator. He also is the Director of Access and Equity. His office is located at 111 Holtzendorrf Hall, 864.656.3181 (voice) or 864.565.0899 (TDD).

Code Plagiarism

All the assignments and projects will be programming related. In the world of Internet, it can be tempting to copy and paste the codes. However, **I take code plagiarism issues very seriously. It is allowed to discuss high-level ideas with classmates. However, copying code or data (either fully or partially) is considered as academic dishonesty. If you use other's code in your assignments or projects, you should put an appropriate reference to it. If you are not sure about the boundary, please contact the instructor**.

Attendance Policy

This course is entirely asynchronous. Successful students will engage with the course material each week and participate on a regular basis. Students should expect to spend 3 hours per credit hour on the course.

Online Conduct

Appropriate online academic conduct means maintaining a safe learning environment based on mutual respect and civility. All participants in Clemson online courses are expected to behave professionally by adhering to these standards of conduct:

- Never transmit or promote content known to be illegal.
- Respect other people's privacy as well as your own.
- Never use harassing, threatening, embarrassing, or abusive language or actions.

Online communication that fails to meet these standards of conduct will be removed from the course. Repeated misconduct may result in being blocked from online discussions, receiving a grade penalty, or being dismissed from the course. Such misconduct in the online environment may also be reported to officials for appropriate action in accordance with University policy. If you ever feel as though our online classroom is inappropriate or uncomfortable, please first contact your instructor with your concerns.

GRADING INFORMATION

The grading system is as follows:

A—Excellent indicates work of a very high character, the highest grade given.

B—Good indicates work that is definitely above average, though not of the highest quality.

C—Fair indicates work of average or medium character.

D—Pass indicates work below average and unsatisfactory, the lowest passing grade.

F—Failed indicates that the student knows so little of the subject that it must be repeated in order that credit can be received.

I—Incomplete indicates that a relatively small part of the semester's work remains undone.

W—Withdrew indicates that the student withdrew from the course or was withdrawn by the instructor after the first two weeks of class work and prior to the last seven weeks of classes, not including the examination period.

Refer to "Grading System" under "Academic Regulations" in the Clemson University's Undergraduate Announcements found on the Registrar's webpage.

Final grades are comprised of the following:

Item	**Percent of Total**
Assignments (10 in total)	30% (3% in each)
Final Term Project	15%
Discussion & Participation	5%
Exam #1	15%
Exam #2	15%
Exam #3	20%
Total:	100%

Final averages between 90.00 and 100 are an A, 80.00 and 89.99 a B, 70.00 and 79.99 a C, 60.00 and 69.99 a D, and below 59.99 an F. Final averages are calculated, and rounded, to two decimal places; i.e., 79.991 is 79.99 and a C, while 79.997 would be rounded to 80.00 and a B.

Assignments (30%)

During the course, you will be given 10 sets of Python programming assignments, which should be conducted *individually*. All the works should be submitted electronically via Canvas (not email attachment) by the due date and time specified by the instructor. Late submissions will receive a score of zero (**no exceptions**). I take code plagiarism issues very seriously. Please refer to the details in the Course Policy section above.

- **It is my policy that getting or giving unauthorized help on individual assignments is a violation of academic honesty, and will lead to an "F" in the course, at a minimum. Don't do it! Get help from me instead!**

- All the assignments will be programming. ***I take code plagiarism issues very seriously. It is allowed to discuss high-level ideas with classmates.*** However, copying code or data (either fully or partially) is considered academic dishonesty. ***If you use other's code in your assignments or projects, you should put an appropriate reference to it.*** If you are not sure about the boundary, please contact the instructor.

Final Term Project (15%)

There will be a term project after the exam #2. The topic will be posted on the course website.

Discussion & Participation (5%)

We all bring experience and knowledge into the classroom, and all class participants should share this and benefit by it. Effective class participation includes:

- **Posting and answering programming questions in the discussion board**
- **Asking questions about concepts from lectures or readings that you agree or disagree with**
- **Sharing your experience or point of view with the class**
- **Building on points raised by others, clarifying issues, or relating topics discussed to previous class discussions.**

Exams (50%)

There will be three online exams. You are responsible for everything that is covered in the classroom, including additional materials that are discussed in class. The exam will consist of multiple-choice questions and programming questions. The exams will be closed book, but a Python reference sheet will be provided before and during the exams. Students are expected to write Python codes based on the specifications.

- **Remote Proctoring Requirements:**
    - All exams will require using **LockDown Browser**. Please review these instructions in Canvas well before the first exam date. Carefully following these guidelines will help minimize technical problems.
    - Having technical problems on an exam as a result of not following directions may not justify obtaining special privileges such as another attempt. If you have followed all the instructions and still encounter technical problems, you must email your instructor as soon as you no longer have access to the exam.
    - **You will need a camera, microphone, and a sufficient Internet Connection** (specifically, upload speed) to take exams using the LockDown Browser software. You must have use of a computer on which you can run the LockDown Browser software.
    - **No software programs or websites other than the LockDown Browser and Canvas** can be running during exams (including email, chats, etc.). No electronic referencing of any kind (e.g., e-book, electronic files/notes, dictionaries, search engines, etc.) is allowed and you cannot use/access any electronic or hard copy resources (e.g., notes, textbook) of any kind (except a reference sheet given by the instructor).
    - **You are not allowed to move away from the Canvas testing interface for any reason and this interface should be the only thing visible on your screen activity.** Use of screen capturing and recording software is strictly prohibited, as well as any other form of copying the exam. Aside from your laptop,
    - Ethernet and power cord (if needed), no electronic devices (e.g., cell phone) of any kind may be present in your testing area. All exams must be completed on an individual basis. No communication, teamwork or collaboration of any kind is allowed.

- It is my policy that getting or giving unauthorized help on exams is a violation of academic honesty, and will lead to an "F" in the course, at a minimum. Don't do it! Get help from me instead!

- The instructor reserves the right to change the letter grade cutoffs.

Tentative Fall 2024 Schedule for DSA 8640

Updated: 08/17/2022

Week #	Date	Topic	Chapter	Assignment
1	Aug. 21	Course Overview		**Install Anaconda**
2	Aug. 26	Python Data Types	Ch. 2	
	Aug. 28	Imperative Programming	Ch. 3	**HW1 Open**
3	Sep. 02	Text Data, Files, and Exceptions	Ch. 4	
	Sep. 04	Execution Control Structure (1)	Ch. 5	**HW2 Open / HW1 Due**
4	Sep. 09	Execution Control Structure (2)	Ch. 5	
	Sep. 11	Execution Control Structure (3)	Ch. 5	**HW3 Open / HW2 Due**
5	Sep. 16	Review, Q&A		**HW3 Due**
	Sep. 18	**Exam #1**		*You will need: fully charged laptop with webcam and microphone, internet connection.*
6	Sep. 23	Containers and Randomness	Ch. 6	
	Sep. 25	Namespaces (1)	Ch. 7	**HW4 Open**
7	Sep. 30	Namespaces (2)	Ch. 7	
	Oct. 02	Object-Oriented Programming (1)	Ch. 8	**HW5 Open / HW4 Due**
8	Oct. 07	Object-Oriented Programming (2)	Ch. 8	
	Oct. 09	Object-Oriented Programming (3)	Ch. 8	**HW6 Open / HW5 Due**

9	Oct. 14	***Fall Break***		
	Oct. 16	Review, Q&A		**HW6 Due**
10	Oct. 21	**Exam #2**		*You will need: fully charged laptop with webcam and microphone, internet connection.*
	Oct. 23	Pandas DataFrame (1)	Jupyter	
11	Oct. 28	Pandas DataFrame (2)	Jupyter	
	Oct. 30	Pandas DataFrame (3)	Jupyter	**HW7 Open**
12	Nov. 04	Pandas DataFrame (4)	Jupyter	
	Nov. 06	Pandas DataFrame (5)	Jupyter	**HW8 Open / HW7 Due**
13	Nov. 11	Web Scrapping (1)	Jupyter	
	Nov. 13	Web Scrapping (2)	Jupyter	**HW9 Open / HW8 Due**
14	Nov. 18	Web Scrapping (3)	Jupyter	**Final Term Project Open**
	Nov. 20	Text Mining (1)	Jupyter	
15	Nov. 25	Text Mining (2)	Jupyter	**HW10 Open / HW9 Due**
	Nov. 27	***Thanksgiving Holidays***		
16	Dec. 02	Text Mining (3)	Jupyter	
	Dec. 04	Review, Q&A		**HW 10 Due / Final Term Project Due (Dec. 8)**
17	Dec. 09 – 13	**Exam #3**		*You will need: fully charged laptop with webcam and microphone, internet connection.*

Note: As the instructor for this course, I reserve the right to adjust this schedule in any way that serves the educational needs of the students enrolled in this course. Students are responsible to be aware of changes announced in class and/or via course website. – Hongki Kim

Kindly check the Academic Calendar for important dates.