Data Science Programming
(2-2-3)

Term: 2023-2024 Fall
Instructor: Dr. Yuan An
Office: Room 1112, 3675 Market Street
Office Hours: Any available time (send me a message requesting for a meeting)
Contact: ya45@drexel.edu
Course Location: 1053
Course Time: MW 12:00pm - 1:20pm

Contacting the Instructor

Email is the best way to contact me. **Be sure to include a subject line, and start the subject with a reference to the course.** For examp le: "(INFO 212) Question" would work. Email without a clear subject may be deleted with spam

Catalog Course Description

Introduces the main tools and ideas in the data scientist's toolbox. Focuses writing interactive and programming code for extracting, cleansing, wrangling, transforming, reshaping, and analyzing data. Covers practical tools and ideas including Linux command line, version control, git, and interactive programming. Studies various Python packages for high performance data analysis.

Curriculum Role

This course is a core course for all Data Science majors and a required course for the Data Science Minor.

Course Rationale

This course prepares students to use practical tools and programming packages to acquire, clean, transform, and analyze various data sets.

Course Outcomes

Upon successful completion of this course, a student will be able to:

- Describe the main steps and key issues in the process of acquiring and preparing data for data analytics.
- Set up data analysis environment by integrating commonly used practical tools and programming packages.
- Explain different types of data storages and formats and apply appropriate tools for extracting and transforming data.
- Create interactive and programming code for cleansing, wrangling, reshaping, visualizing and analyzing various data sets.
- Explain the concepts of aggregation and grouping, and apply tools and write programs to aggregate and group data.

Prerequisites

INFO 153 or CS 172

Recommended Textbook

1. *Python for Data Analysis* by Wes McKinney. Publisher: O'Reilly Media.

References

1. Pandas User Guide: https://pandas.pydata.org/docs/user_guide/index.html
2. Numpy User Guide: https://numpy.org/doc/stable/user/
3. Scipy Reference Guide: https://docs.scipy.org/doc/scipy/reference/
4. Matplotlib User Guide: https://matplotlib.org/stable/users/index.html
5. Seaborn User Guide: https://seaborn.pydata.org/tutorial.html

Additional References

TBD

Tentative Schedule

The schedule is tentative and is likely to vary somewhat based on knowledge of students in the class section. The table below shows the initial schedule for the term.

Week	Topics	Reading	Tasks
1	• Setup • Review of Python Basics • IPython • Jupyter Notebooks	Chaptes 1 and 2 (1)	Assignment 1 out
2	• Built-in Data Structures • Functions and Files	Chapter 3 (1)	
3	• Numpy Basics • Arrays • Vectorized Computation	Chapter 4 (1)	
4	• Pandas Basics • DataFrames • Indexing	Chapter 5 (1)	Assignment 1 Due Assignment 2 out
5	• Matplotlib Basics • Plotting and Visualization	Chapter 9 (1)	
6	• Data Loading • Data Storage • Data Formats • Use Command line Tools and Git	Chapter 6 (1)	

7	• Data Cleaning • Data Preparation	Chapter 7 (1)	Project proposal out
8	• Data Wrangling • Join, Combine, and Reshape	Chapter 8 (1)	Assignment 2 due Assignment 3 out
9	• Data Aggregation • Group Operations	Chapter 10 (1)	Project report out
10	• Advanced Topics • Time Series • Data Analysis Examples	Chapters 11-14 (1)	
11			Final Project Due Assignment 3 due

Grading

Your course grade will be based on: **course participation, homework assignments, exam, and pair projects**. The grade is computed as follows:

Task	**Weight**	**Due**
Course participation and labs	40%	
Assignment 1	10%	Week 4
Assignment 2	10%	Week 8
Assignment 3	20%	Week 11
Final Project Proposal	5%	Week 9
Final Project	15%	End of the term

All work is graded on a numeric scale of 0 to 100. Assignments are graded based on their content (whether you answered the questions correctly and well) and their form (did you produce a submission that has a professional appearance that follows course requirements). A single grade is computed for the assignment based on those two factors.

Conversion from points to letters is given in the following table:

>=99	A+	77-79	C+
93-98	A	73-76	C
90-92	A-	70-72	C-
87-89	B+	67-69	D+
83-86	B	63-66	D
80-82	B-	<=62	F

Exam and Project

The midterm exam is canceled. The weight of the midterm is factored into the labs after Week 5. Therefore, in-class labs will be graded for completion and correctness after week 5 in addition to attendance.

Final Project

The final projects must be proposed and developed by student groups each of which consisting of up to 3 students. Each team must obtain an approval from the instructor for their proposed project. Unapproved projects will be graded as zero when weighed in the final grade.

Submitting Assignments

1. You must submit **assignments** electronically via Blackboard Learn **no later than 11:59pm** on the due day. The **name(s)** and **student ID(s)** of all submitters of the assignment along with the **course number** and **assignment number** must be clearly printed on the first page.
2. University rules and policies regarding academic honesty are followed to the letter.

Late or Missing Work

Each item of course work you turn in has a due date and time:

- Deliverables are due at the time and date indicated by instructions. Written assignments are due no later than **11:59pm** on the due day.

- There will be a 10% (absolute value) deduction for each day of lateness, to a maximum of 3 days; assignments will not be accepted beyond that point. Missing work will earn a zero grade.

Re-Marking

If you are dissatisfied with a grade or point deduction, you can request re-marking. All re-marking requests must be done through written (paper or email) descriptions of why you think the grade is in error. Please note that it is very rare that changing disputed grades on assignments or exams actually affects a student's calculated letter course grade. The corrections are usually insignificant. If you wish to appeal a course grade, school policies apply.

Academic Honesty

The Drexel University Academic Honesty Rules and Procedures (as stated in the student handbook) will be adhered to strictly. Students who commit plagiarism or cheat on assignments may receive an F grade for both the assignment and the course.

In order to avoid plagiarizing material, observe the following:

If you work on an assignment with another student or a group of students, be certain that your final, individual paper is your own work or the work of your project group (for group projects) unless otherwise specified by the professor. While you might want to discuss the assignment with other students, you must, in your paper, express your own ideas in your own way.

If you use printed or electronic resources in your papers, be sure to attribute the sources you have used. This can be done by quoting the material or by paraphrasing the material and, in either case, listing the source in an annotated bibliography. Use standard notation when citing references.

The college is requiring students to append a statement to deliverables (for example: papers, projects, exams) indicating that the work submitted is their own. Deliverables will not be graded without a separate certification page.

Please sign the following academic honesty statements by typing your name and date, and submit it through the Blackboard system as an individual assignment.

==========

I certify that my work in this course will be entirely my own work. I will not quote the words of any other person from a printed source or a website without indicating what has been quoted and providing an appropriate citation. I will not submit my work in this course to satisfy the requirements of any other course.

Name/Signature_____
Date _____
==========

Class Attendance and Participation

You will get the most out of this course if you come to class on time, participate and conduct yourself in a professional manner. Skipping classes will hinder your ability to do well in the course and to learn this important subject. You will also miss handouts, assignments, and announcements. Participation in class is important; there are no stupid questions or comments.

Class attendance is expected. Attendance includes arriving when the class is scheduled to start and staying for the duration of the class period. Roll may be taken in this course and you are expected to acknowledge your attendance at least 90% of the time that attendance is taken in the course**. I reserve the right to lower any earned course grade if a student fails to meet attendance requirements.**

Disabilities

Accommodation of Special Needs - Students with disabilities requesting accommodations and services at Drexel University need to present a current accommodation verification letter (AVL) to faculty before accommodations can be made. AVL's are issued by the Office of Disability Services (ODS). For additional information, contact ODS at 3201 Arch St., Street, Suite 210, Philadelphia, PA 19104, 215.895.1401 (V), or 215.895.2299 (TTY). http://www.drexel.edu/oed/disabilityResources/students/

Withdrawal of the Course

For dropping or withdrawing from the course, please refer to the university policies at:
http://www.drexel.edu/provost/policies/course_drop.asp
http://drexel.edu/provost/policies/course-withdrawal/

Class Cancellation

On rare occasions, instructors may be delayed or unable to attend a scheduled class due to unforeseen circumstances. In the event that an instructor does not appear in class and has not

notified the class of his/her expected arrival time, class is cancelled 15 minutes after the scheduled start of class. More information about class cancellations can be found at http://drexel.edu/provost/policies/cancellation_instructor_absence/

Class Lecture Recording
Lectures and class discussions may be audio-recorded and streamed or rebroadcast for educational purposes only.

Incomplete Policy
Incomplete grades are contingent upon instructor approval and will only be considered in extenuating circumstances beyond the student's control. The instructor is under no obligation to offer an incomplete grade. At least 80% of the graded coursework must have already been completed in order for an incomplete grade to be considered (per the recommendation of the Provost's Office). An incomplete contract with an instructor-determined due date for delivery of the completed work must be completed by the student and the instructor. It can be found here: http://www.drexel.edu/provost/policies/pdf/forms/incomplete.pdf.

Support for Equality and Diversity
Drexel University strives to promote an environment of equality of opportunity and compliance with University policies and federal, state and local laws prohibiting discrimination based upon race, color, religion, gender (sex), marital status, pregnancy, national origin, age, disability and veteran status. Students, faculty, and staff with questions about or complaints concerning discrimination, harassment, and/or retaliation should contact the Office of Equality and Diversity at (215) 895-1403 or http://www.drexel.edu/oed/

Student Conduct and Community Standards
Drexel University expects that all students as well as student organizations will conduct themselves responsibly and in a manner that reflects favorably upon themselves and the University. Check out here:
http://www.drexel.edu/studentlife/community_standards/overview/ for the university policies, rules, regulations, and standards of conduct.

Strategies for Success
- Attend class regularly and complete all assignments on time.
- Complete the assigned reading for each week before the class of that week.
- Take good notes in class.

Course Evaluation
Your feedback about the course and instructor is the only way instructors and academic units can improve the quality of a course and its content. Courses administered by the College of Computing and Informatics are evaluated electronically via AEFIS. Students will receive all necessary information via email by the 8th week of classes (or by the 4th week of classes in case of accelerated courses). The evaluations are entirely confidential and will preserve your anonymity.

Syllabus Changes
The instructor reserves the right to make changes to this syllabus if circumstances warrant such change. All changes will be provided to students in writing.

Notice: Appropriate Use of Course Materials

It is important to recognize that some or all of the course materials provided to you may be the intellectual property of Drexel University, the course instructor, or others. Use of this intellectual property is governed by Drexel University policies, including the policy found here: ***https://drexel.edu/it/about/policies/policies/01-Acceptable-Use/***

Briefly, this policy states that course materials, including recordings, provided by the course instructor may not be copied, reproduced, distributed or re-posted. Doing so may be considered a breach of this policy and will be investigated and addressed as possible academic dishonesty, among other potential violations. Improper use of such materials may also constitute a violation of the University's Code of Conduct found here: ***https://drexel.edu/cpo/policies/cpo-1/*** *and will be investigated as such.*