IDS 400: Programming for Data Science in Business (Fall 2023)

Instructor: Aida Sanatizadeh

Teaching Assistant: Tirunagari Dharma (Lab/OH: Mondays 10-11 AM online on zoom)

Lectures: Thursdays 6:30:00 pm-9 pm (08/24/2023-12/07/2023)

Contact Instructors: Piazza (you can log into Piazza through Blackboard and send a private message to ask your TA or me questions or to make an induvial appointment)

1. **Course Description**

Data volume has grown exponentially in recent years due to the development of technology. Companies are also increasingly employing data analytics techniques to leverage the benefit of large-scale datasets. It has already been well-recognized that companies can significantly improve their profitability through data analytics. This class, IDS 400, Programming for Data Science in Business is an introductory class to help you learn the basics of programming for data analytics, and how to program in Python to solve real-world problems. This will be very helpful for your jobs in the future. This course will first cover the basics of data science, then move to foundations in Python programming, and gradually to data analysis using Python and external packages.

The (tentative) list of topics we will cover in this course is:

- Basic programming concepts
- Install and run Python programs.
- Variables and expressions
- Control structures: sequence, branch, iteration
- List, tuple, and dictionary in Python
- Functions, parameters, and recursion in Python
- String operations
- File operations and modules
- Visualization
- Data manipulation: Pandas
- Optimization and statistics in Python: SciPy
- Machine learning package: Mlpy, scikit-learn.

While this course does not require any prior programming knowledge of Python, students enrolled in the class are expected to have some familiarity with programming in other languages, at the introductory level (Such as Java, R, etc.).

Learning Outcomes:
By the end of this class, students are expected to:

1. Understand the basics of data science in programming.
2. Understand the basics of Python, such as syntax, semantics, iterative control, data structure, etc.
3. Understand how to use existing Python packages.
4. Apply Python package for data storage, manipulation, and cleaning.
5. Understand how to perform machine learning in Python.
6. Be able to use Python to solve real-world problems.

2. **Textbooks**:

While we do not have any required textbook for the course, here are a few recommended books/online resources to learn Python:

- Think Python—How to Think Like a Computer Scientist, by Allen B. Downey,

- (1st edition for Python 2, 2nd edition for Python 3). You can download it here or read it in HTML.

- Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython, 2nd Edition, by Wes McKinney, O'Reilly Media, 2017. The E-book is available at the UIC library website.

- Python Data Visualization Cookbook. This is a book for those who want to explore more on visualization in Python.

- MIT Open Courseware: A Gentle Introduction to Programming Using Python Google's Python class.

3. **Software**

We will mainly use Jupyter Notebook as the programming environment. You can download it through Anaconda (https://www.anaconda.com/download).

4. **Attendance**

I strongly encourage you to attend all lectures. There might be some random attendance checks. I will deduct 5 points if you are absent for all checks. Failing to attend exams or final projects will receive an "F" for the course except for extreme reasons.

5. **Piazza**

We will use the Piazza (you can log into this system through Blackboard) for a discussion board and for posting updates. Participation on the discussion board counts towards your grade. A constructive, relevant, non-redundant post earns a Piazza point.

Please sign in to Piazza by clicking on the following link (or through Blackboard) and read the **welcome to Piazza**. Here are some tips and guidance:

- Please provide enough details in the questions that you post. Your questions should not be superficial. Before posting your questions, please search on Google first.

- I highly encourage students to answer each other's questions. You can also find teammates in Piazza for the final project.

- You can post questions anonymously/non-anonymously.

- Students can send me a private message in Piazza if you cannot post it to the entire class.

6. Lab

We will have a total of 10 labs. Your TA will solve the lab questions and you may ask your questions about either the labs or assignments during the lab time. I highly recommend trying to solve the lab questions beforehand.

Students are required to submit at least 6 labs. Failure to do so will automatically fail this course. Of course, you are encouraged to do all the labs and I will consider this as a bonus point.
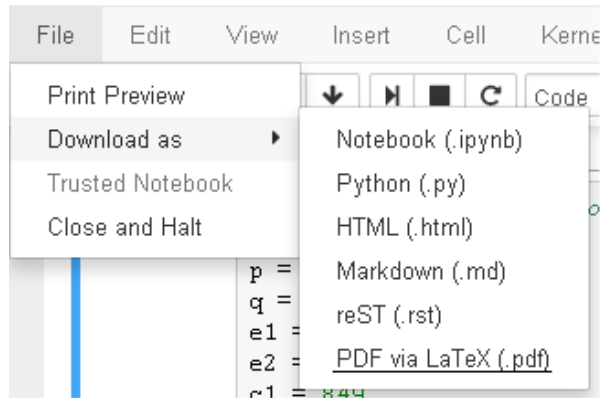
You will need to submit each lab before the next lecture starts.

7. Assignments

There are 6 to 7 assignments across this semester. For each assignment, you will submit two files (IPYNB and PDF) following the format below:

- Lastname_firstname_assignment1.pdf

- Lastname_firstname_assignment1.ipynb

Please note that both files **should show the output of your code** as well. You can save the PDF of your code following the steps shown in the image below.

File   Edit   View   Insert   Cell   Kerne

Print Preview

Download as   ▸   Notebook (.ipynb)

Trusted Notebook   Python (.py)

Close and Halt   HTML (.html)

Markdown (.md)

reST (.rst)

PDF via LaTeX (.pdf)

```
p =
q =
e1 =
e2 =
c1 = 844
```

While you are allowed to submit your assignment late, you will receive a 10% grade penalty for every additional day after the deadline. As this is a generous late submission policy, penalties are strictly enforced, and I will not grant extensions. Please plan accordingly, and do not leave submission for the last minute.

- You should do all assignments and exams by yourself without copying from online resources or your classmates or using ChatGPT. Detection of plagiarism will typically result in a **FULL NEGATIVE GRADE (-100)** for the assignment. Note that a score of -100 may not be dropped from your record.

- You will be allowed to take a make-up exam only in cases of medical or personal emergencies, which must be verified. If such an emergency occurs, call me, or leave a note as soon as possible. Non-emergencies include weddings, conferences, etc.

8. **Final Project**

You can only learn programming through coding by yourself. That is why I emphasize a lot on the final project. You can form a group of up to four students to do the project. The deliverables of the course project include a project report, Python code plus a class presentation. The presentation constitutes 50% of the grade and the project report and Python code deliverables are 50%. The presentation is graded based on content, structure, design, and delivery. By content, we mean that the students have been able to include all the necessary pieces of information needed to achieve the objective. By structure, we mean that the students have included in the presentation introductory, main body, and conclusion slides which follow each other in logical succession. By design, we mean that the students have been able to use layouts, graphics, charts, videos, and other visual aids effectively and enhance their understanding of the topic presented. The second deliverable of the project, the project report, and Python code, is graded for quality and completeness (whether this application achieves the goal you claim).

If any one of the team members requests a peer evaluation within the team, I will conduct the evaluation and adjust the project grade accordingly.

9. Grading Criteria

A (> 90); B (80-89); C (70-79); D (60-69); F (< 60).

Component	Weight
Attendance	10%
Assignment	50%
Quiz	20%
Final project	20%

10. Policy for Missed or Late Work

Late submission will receive credit deduction:

0-1 day	10%
1-2 days	20%
2-3 days	30%
3-5 days	50%
>5 days	will not accept

11. Academic Integrity Policy

If you are struggling in the course please ask for help from us in the instructional staff, rather than representing someone else's work as your own. We believe you are capable of learning the concepts in the course and doing the work, and are committed to helping you get there. We expect you to treat other class members with respect, making a genuine effort on all assignments, discussions and class activities. Submitted work must be your own unless the assignment explicitly is intended to be done in a group.

Note that you are guilty of academic dishonesty if you extend or receive any kind of unauthorized assistance. Absolutely no transfer of program code between students is permitted (paper or electronic), and you may not solicit code from family, friends, or online forums. Other examples of academic dishonesty include emailing or showing your program to another student, copying-pasting code from the internet, working in a group on a homework assignment, and allowing a tutor, TA, or another individual to write an answer for you. Academic dishonesty is unacceptable, and penalties range from failure to expulsion from the university; cases are handled via the official student conduct process.