DAMG 6105: Data Science Engineering with Python

Course Information
Course Title: Data Science Engineering with Python
Course Number:DAMG 6105
Term and Year: Fall 2024
Credit Hour: 4 credits
CRN: 14796
Course Format: On the ground

Instructor Information
Full Name: Ramkumar Hariharan
Email Address: rmhariharan@northeastern.edu
Office Hours: TBD

Instructor Biography
Ramkumar Hariharan is Program Director and Data Science Engineering Faculty at Northeastern, Seattle. He is also Senior Scientist at Northeastern's Institute for Experiential AI (EAI). Previously, he has led multiple high-impact data-driven projects at some of the leading institutes in Seattle. These include Fred Hutch, University of Washington (UW), and the Institute for Systems Biology. His areas of focus include data analyses, data visualization, and predictive analytics of both structured and unstructured data.Ram has a 17-year history of developing and delivering more than 20 computational, biomedical, and data science courses at a variety of levels. His courses, lectures, online teaching, and motivational talks have been overwhelmingly well-received in Seattle, Japan and in India. Ram has been on Television in India and in the US. Ram serves as affiliate of UW e-sciences institute, bootcamp leader at General Assembly, and mentor with Springboard. He has also led education and training programs for Fred Hutch. He specializes in using powerful, yet simple analogies to explain seemingly complex computational and data science concepts and math.Ram's teaching philosophy is grounded in one strong belief: there is no one size fits all approach to teach, or to learn a new concept.

Teaching Assistant Information
Full Name: Zhendong Yang
Email Address:yang.zhend@northeastern.edu
Office Hours:TBD

Course Prerequisites
Please check academic catalog
Course Description Are you looking to explore a high-paying career as a data scientist in the future? Or you are thinking about how best you can hone your existing skills towards becoming

a professional data scientist? If you are then, check out DAMG 6105. Since we have a very different teaching philosophy, individuals from a variety of backgrounds and strengths will find this course useful. This means that if you currently have just high-school level math knowledge, you can still think about taking this course and aspire to become a good data scientist. The goal of this course is to introduce the fundamentals of machine learning (ML) and data science with Python. This course is aimed at individuals who are looking to learn and apply state-of-the-art ML algorithms to real-world datasets using the Python programming Language. There are two unique and foundational pillars of this course— (1) We use real-world situational analogies to intuitively understand the algorithmic aspects of data science and ML, and (2) We use real-world datasets and use powerful machine learning libraries in Python, to put these algorithms to practice.

Course Learning Outcomes

1) Proficiency in Python Data Science Fundamentals: Students will attain proficiency in fundamental concepts of data science using Python, including numerical analysis, scientific computing, and manipulation of large datasets.
2) Data Analysis and Manipulation: Students will develop advanced skills in data manipulation and analysis using Python libraries be able to support large and high-dimensional datasets, performing group and pivot operations, and managing missing data efficiently.
3) Effective Data Visualization Techniques: Students will master data visualization techniques using Python libraries to create visually appealing and informative plots to communicate insights effectively from data analysis.

Required Tools and Course Textbooks

1. Head First Python: A Brain-Friendly Guide, Paul Barry, O'Reilly, 2nd or New Edition
2. Python Data Science Handbook: Essential Tools for Working with Data, Jake VanderPlas, O'Reilly, 2nd Edition
3. Python for Data Analysis : Data Wrangling with pandas Numpy, Wes Mckinney, O'Reilly, 3rd Edition

Course Schedule/Topics Covered.

Week	Date	In Class Topic	Assignment Due
1	09/10	Course admin trivia, Why even Python it ?	
2	09/17	Variables (Scalars, Lists, Tuples, Dicts)	Problem set I released
3	09/24	Flow of Control & Modular Python	Project Title, Theme Due
4	10/01	Wrangling Time : pandas I	Problem set I Due
5	10/08	Wrangling Time : pandas II	
6	10/15	Machine learning Foundations	Problem Set II Released
7	10/22	K- Nearest neighbors I	
8	10/29	K – Nearest neighbors II	Problem Set II Due
9	11/05	K – Nearest neighbors III & ML Evals	

10	**11/12**	**Midway Break**	
11	**11/19**	**Trees to Decide**	**Problem Set III Released**
12	**11/26**	**From Trees to forests**	
13	**12/03**	**Random Forests in Python**	**Problem Set III Due, Project Submission Due**
14	**12/10**	**Final Project Presentation**	
15	**12/17**		
16	**12/24**		

Assignment Grading

Please insert all assignment grades and weights for the course. Example below:

- Assignment 1 – 25%
- Assignment 2 – 25%
- Assignment 3 – 25%
- Final Project – 25%

Grading Scale

	87-89.9% B+	77-79.9% C+	
95-100% A	84-86.9% B	74-76.9% C	
90-94.9% A-	80-83.9%B-	70-73.9% C-	69.9% or below F

Attendance/Late Work Policy

Attendance Policy

Students registered in MGEN courses (INFO, CSYE, and DAMG) are allowed **a maximum of 2 absences per course, with 3 or more absences resulting in an automatic 'F' for that course.** Students are expected to inform their instructors of any absences in advance of the class; if a student is sick long-term or experiences a medical issue that prevents class attendance, it is strongly encouraged that they speak with their Academic Advisor (coe-mgen-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. Should a student anticipate being unable to attend 3 or more classes, they should discuss their situation with their Academic Advisor to explore other types of leave in accordance with the University's academic and global entry expectations. International students should review the Office of Global Services webpage to understand their visa compliance requirements.

Teaching Assistants (TAs) or Instructional Assistants (IAs) will be present at each class to collect student attendance.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus. Students must communicate with the faculty prior to the deadline if they anticipate work will be submitted late. Work submitted late without prior communication with faculty will not be graded.

End-of-Course Evaluation Surveys
Your feedback regarding your educational experience in this class is particularly important to the College of Engineering. Your comments will make a difference in the future planning and presentation of our curriculum.

At the end of this course, please take the time to complete the evaluation survey at **https://neu.evaluationkit.com**. Your survey responses are **completely anonymous and confidential**. For courses 6 weeks in length or shorter, surveys will be open one week prior to the end of the courses; for courses greater than 6 weeks in length, surveys will be open for two weeks. An email will be sent to your Northeastern University Mail account notifying you when surveys are available.

Academic Integrity
A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to http://www.northeastern.edu/osccr/academic-integrity-policy/ to access the full academic integrity policy.

MGEN Student Feedback
Students who would like to provide the MGEN unit with anonymous feedback on this particular course, Teaching Assistants, Instructional Assistants, professors, or to provide general feedback regarding their program, may do so using this survey: https://neu.co1.qualtrics.com/jfe/form/SV_cTlAbH7ZRaaw0Ki

University Health and Counseling Services
As a student enrolled in this course, you are fully responsible for assignments, work, and course materials as outlined in this syllabus and in the classroom. Over the course of the semester if you experience any health issues, please contact UHCS.

For more information, visit https://www.northeastern.edu/uhcs.

Student Accommodations
Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit https://drc.sites.northeastern.edu.

Library Services
The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and for education specific resources, visit https://library.northeastern.edu
Network Campus Library Services: Northeastern University Library Global Campus Portals

24/7 Canvas Technical Help
For immediate technical support for Canvas, call 617-373-4357 or email help@northeastern.edu

Canvas Student Resources: https://canvas.northeastern.edu/student-resources/

For assistance with my Northeastern e-mail, and basic technical support:
Visit ITS at https://its.northeastern.edu
Email: help@northeastern.edu
ITS Customer Service Desk: 617-373-4357

Diversity and Inclusion
Northeastern University is committed to equal opportunity, affirmative action, diversity, and social justice while building a climate of inclusion on and beyond campus. In the classroom, members of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration, and an awareness of global perspectives on social justice.

Please visit http://www.northeastern.edu/oidi/ for complete information on Diversity and Inclusion

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, transgender students, faculty, and staff.

In case of an emergency, please call 911.

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