Introduction to Scientific Computing (Data Science with Python) I

DATA 1401
Unique Number: 28122
Spring 2021

Instructor:	Amir Shahmoradi
office:	SEIR 365
e-mail:	a.shahmoradi@uta.edu
e-mail:	shahmoradi@utexas.edu
e-mail:	a.shahmoradi@gmail.com
Office/Lab/Help hours:	**dataCAVE** (computer lab in the UTA Main Library) Online/in-person by appointment

Class start/end:	January 20, 2021 – May 3, 2021
Lecture meeting times:	Monday – Wednesday 2:30 – 3:50 pm
Lecture meeting place:	Microsoft Teams / In person with prior notice
Lab meeting times:	Friday 1:00 – 2:50 pm
Lab meeting place:	Microsoft Teams / In person with prior notice

Teaching Assistants:	Anoop Kunjumon
office:	Microsoft Teams (same room as class)
e-mail:	anoopkunjumon.scariah@mavs.uta.edu
office hours:	Monday – Wednesday 9-11 am

COVID-19 Notes:

Given the current global pandemic situation, in-person meetings will be limited. All lectures, HW, and quizzes can be also done and submitted online with no contact with the instructor. Students will be allowed to attend the classes and submit their HW/Quizzes completely online if they are concerned about their safety and health or the health of their family members. In such case, please inform the instructor of your intention to attend the classes online.

COURSE OBJECTIVES / ACADEMIC LEARNING GOALS

This is the first of a two-course sequence providing the necessary foundations in scientific computing for Data Science majors. It introduces a number of operating systems, languages, and tools using examples and contexts from natural and behavioral sciences. **Prerequisite**: DATA 1401, **or** permission of the instructor.

The primary objective of this course is to learn basic computer programming concepts and apply them to Data Science related problems. By the end of the course, you should have a good understanding of general programming foundations and practices and be able to analyze data-intensive problems and develop computational solutions for them, collaboratively within a team. We will achieve this by learning how to program in popular operating system environments such

as Linux using popular high-level programming languages such as Python. Although not essential, some prior level of familiarity with programming concepts is desirable for this course.

COURSE SCHEDULE

The following is a tentative outline of topics to be covered:

- Version Control Systems (VCS): Principles of professional project management and collaborative programming with the use of Git.
- Programming History – Operating Systems – Round-off and Truncation Errors – Number Systems – Introduction to Cloud Computing Resources.
- Brief Introduction to Principles of Programming using Python as the main language: Python development environment – general installation guidance – available editors – syntax rules – variables and data types – conditionals – looping – input/output – functions – Exception Handling – Object Oriented Programming – Array Computing and Performance Optimization – Wrappers, and Cross-language Interoperability.
- Applications of Programming in Data Science: Matrix Operations – Data Manipulations – plotting – symbolic calculations – libraries (e.g., Numpy, Scipy, Pandas, scikit-learn, mlpy, seaborn, …).

COURSE TEXTBOOKS

No textbook is required for this course. Online class lecture notes will be used as reference. However, a list of textbooks for those who are interested to self-educate themselves or go beyond class syllabus will be provided on the first day of the class.

COURSE LOGISTICS

Grading:
Weekly Homework: 25% (Assignments might not be weighted equally)
Weekly Quizzes: 25%
Midterm Exam: 25%
Final Project: 25%

Homework Policy:
There will be approximately one homework per week. Assignments will be due before lecture begins and should be added to an online repository determined by the instructor. No late assignments will be accepted. No exceptions to the homework policy will be made without prior instructor approval.

Examinations:
There will be one midterm exam, but no final exam. Students will have to complete a project in place of the final exam, in collaboration with their teammates who are determined randomly after the midterm.

Attendance:

Regular attendance is expected. Any absence requires prior approval from the instructor, or compelling evidence of illness or an official letter from the university administration. Student attendance will be randomly checked.

Scholastic dishonesty: All students are responsible for upholding the University rules on scholastic dishonesty. Students who violate University rules on scholastic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on scholastic dishonesty will be strictly enforced.

Other matters: The University of Texas at Arlington provides, upon request, appropriate academic adjustments for qualified students with disabilities. Any student with a documented disability (physical or cognitive) who requires academic accommodations should contact the UTA's Office for Students with Disabilities as soon as possible to request an official letter outlining authorized accommodations. For visit https://www.uta.edu/disability/.

Your Expectations:

For the current offering of this course, we will cover the principles of computer programming using Python programming language. Specifically, upon completion of this course you will be familiar with,

- programming paradigms,
- principles of software maintenance and collaborative project development,
- differences between compiled and interpreted programming languages,
- how to use Python as a simple calculator,
- how to use Python as an advanced scientific computation and graphics toolbox,
- how to use Python to collect, clean, and analyze data and make a scientific inference,
- how to formulate and cast a scientific problem in the form of a sequence of computational algorithms.