Data 3402: Python for Data Science 2

Spring 2026

Instructor Information

Instructor
Amir Farbin

Office Number
CPB-340

Office Telephone Number
817-272-1178

Email Address
afarbin@uta.edu

Faculty Profile
https://mentis.uta.edu/explore/profile/amir-farbin

Office Hours
Best time to quickly chat with me is after class. Otherwise, I am generally available and am happy to meet virtually. If you like to speak to me, please send a chat message via Teams, letting me know when you would be available in the next 24-48 hours and I will follow up.

Course Information

Section Information
DATA 3402

Time and Place of Class Meetings
- Lectures: MW 1-2:20 in UH 10
- Lab: F 1-2:50 Virtual

Description of Course Content
DATA 3401 -- Python for Data Science 1
This is the first of a two-course sequence offering the foundations of Python programming in the context of data science. It introduces the full syntax of the Python language as it overviews structured, functional, and object-oriented programming methodologies. It also provides a basic conceptual understanding of computing and introduces Unix command-line tools, software employed in data science, such as git and Jupyter, and Python libraries such as numpy, matplotlib, and Pandas.

DATA 3402 -- Python for Data Science 2
This is the second of a two-course sequence offering the foundations of Python programming in the context of data science. It reinforces concepts presented in DATA 3401 with greater depth and a focus on application to various problems in data science, while further exploring the python library ecosystem.

Student Learning Outcomes
Upon successful completion of this course, students will be able to:

1. Design and implement computational solutions using structured, functional, and object-oriented programming paradigms.
2. Construct simplified prototype implementations of core data science abstractions (e.g., arrays, data containers, histograms, and plotting systems) in order to understand their underlying design.
3. Use professional data science libraries (e.g., NumPy, Pandas, Matplotlib, scikit-learn) as informed tools rather than black boxes, articulating the design motivations and tradeoffs behind their use.
4. Develop and analyze computational approaches to open-ended, real-world problems using simulation, data processing, and statistical reasoning.
5. Build, evaluate, and interpret basic machine learning models for classification and regression within well-structured computational workflows.
6. Transition from prototype implementations to scalable, library-based solutions, evaluating correctness, performance, and clarity.

Required Textbooks and Other Course Materials
No text book required for the course. All material will be made available on GitHub for the students.

Lectures
The course lectures will be in-person, meaning you are expected to be in class. Unless I'm sick, need to be quarantined, or traveling, I will be teaching in the class room. Nonetheless, I will also use Teams to record the lectures and enable students to connect remotely if necessary. We will take attendance in every class period including labs. Attendance is 20% of your grade. If you are missing class for any reason, you need to write the professor and the TAs and let them know.

Classroom/Lecture Recording Policy
Faculty maintain the academic right to determine whether students are permitted to record classroom and online lectures. Recordings of classroom lectures, if permitted by the instructor or pursuant to an ADA accommodation, may only be used for academic purposes related to the specific course. They may not be used for commercial purposes or shared with non-course participants except in connection with a legal proceeding.

All session of this course will be recorded on Teams and available to students to review.

Labs
The course labs will be all virtual and run via Teams. I've found virtual labs to be much more effective than in person.

Course Communications
All course communications will be made on Teams. Please do not use e-mail. Also make sure you have Teams app installed and you are looking at Teams notifications. I also encourage students to ask questions and interact with me and the TA via Teams chat.

Descriptions of major assignments and examinations
Your grade will primarily based on your performance on weekly assignments (labs) and your final project. See grading breakdown below. About half-way through the course, you will be required to take the python proficiency exam from the python institute (pythoninstitute.org/pcep) at a cost of $59. Your performance on the exam will not affect your grade in the course, but submission of your results is 10% of your grade.

Technology Requirements
You have several options on how to work on assignments. In all cases you will need access to a computer (laptop or desktop) with a keyboard and mouse. Tablets are not recommended. Your options:

- **Laptop (Preferred):**
  - OS: Windows (with WSL installed), MacOS, or Linux
  - You will install and run everything locally. For the most part, almost any laptop will do. Later in the course, some of the assignments may require significant disk space / memory. If your laptop can't handle it, you can switch to use cloud option (see below).
  - You should bring your laptop to the lab sessions.
  - Note: A limited number of laptops (5) are available for checkout, specifically for this course, at the library.
- **Desktop:**
  - OS and software setup will be identical to laptop option.
  - For in-person labs, you will have to fall back to cloud option below.
- **Cloud**:
  - You have the option of working on labs/projects on Google's Colab platform and storing your files in Google Drive.

You are required to either use Linux, MacOS, or Windows with WSL.

Grading Information

Grading

Student performance in this course will be evaluated using the following components:

Attendance and Quizzes — 20%

Attendance is required for all lectures and labs. Quizzes will be used primarily to encourage attendance, assess student progress, and initiate focused discussion. Quizzes are generously graded and are intended to support learning and help boost overall grades.

Students are permitted up to two (2) unexcused absences during the semester.

Python Proficiency Exam — 10%

Midway through the course, students are required to complete the Python Proficiency Exam. Performance on the exam will not affect the exam score itself; however, completion and submission of results is required to receive credit.

Laboratory Assignments — 50%

Approximately eight (8) laboratory assignments will be given throughout the semester. Each lab typically spans one to two weeks.

To account for unforeseen circumstances, the lowest lab grade will be dropped, including absences due to illness, unless prior arrangements have been made with the instructor.

Final Project — 20%

The final project will take the form of a Kaggle-style data science challenge. Students will present their work during the final exam period.

All grades will be curved, if need be. The exact curving methodology will be the topic of a lecture. It is extremely important to not fall behind in this course.

Expectations for Out-of-Class Study
You are expected to spend about 10 hours per week working on this course outside of lecture and lab hours.

Help
In addition to the instructor, the course TA will be available via Teams for help.

Course Schedule

The following schedule and topics list is tentative. This course is continuously under development and I will adjust course content to the needs of students. *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.*

Tentative Course Schedule and Topics

Module 1 — Course Orientation and Computing Foundations (LO1, LO6)
- Course introduction and expectations
- Instructor research overview
- How computers work
- From transistor to modern computing systems
- Storage and filesystems
- Firmware → Operating System → Applications
- Machine language → Python

Lab 1: Development environment setup
- Linux / WSL / Google Colab
- GitHub workflow
- Basic shell commands

Module 2 — Structured Programming and Problem Decomposition (LO1, LO4)
- Review of structured programming
- Decomposing problems into logical components
- Building a simple game (Checkers)

Lab 2: Tic-Tac-Toe implementation

Module 3 — Functional Programming and Stochastic Simulation (LO1, LO2, LO4)
- Functional programming concepts
- Probability and statistics
- Random numbers and fundamental distributions
- Histograms and data summarization
- List comprehensions and functional tools (functools, etc.)
- Data processing workflows

Lab 3: Random processes, distributions, histograms, and Monte Carlo simulation

Module 4 — Object-Oriented Programming and Data Abstractions (LO1, LO2)
- Object-oriented programming concepts
- Design patterns and UML
- How object-oriented systems work in practice

Module 5 — Numerical Computation and Visualization Systems (LO2, LO3, LO6)

- Tensor operations and array semantics
- Introduction to numerical computation (NumPy)
- Plotting systems and visualization design
- Canvases, coordinate systems, rendering

Lab 4: Building and using plotting software (Matplotlib)

Module 6 — Data Representation, Persistence, and Software Design (LO2, LO3, LO6)
- Data persistence and representation
- CSV readers and tabular data
- DataFrame abstractions (Pandas)
- Scripting vs. building software systems
- Case studies: gradebook and simulation systems

Lab 5: Data representation and persistence (Pandas)
Lab 6: Blackjack simulation (LO1, LO4, LO6)

Module 7 — Data Analysis and Statistical Reasoning (LO3, LO4)
- Exploratory data analysis
- Visualization for insight
- Data processing and summarization
- Hypothesis testing

Module 8 — Domain Case Study: High Energy Physics(LO4, LO6)
- Domain introduction: High Energy Physics
- Scientific data pipelines and challenges

Lab 7: High Energy Physics data analysis

Module 9 — Machine Learning Foundations (LO4, LO5, LO6)
- Introduction to machine learning
- Supervised learning
- Classification
- Regression
- Model evaluation and interpretation

Lab 8: Classification task

Module 10 — Projects and Advanced Topics (LO4, LO5, LO6)
- Project proposals
- Ongoing project check-ins during lectures
- Advanced and targeted topics

Module 11 — Looking Forward: Deep Learning and Computation (LO3, LO6)
- Deep learning overview
- Computational scaling
- Multithreading and multiprocessing
- TensorFlow and PyTorch as computation engines

Lab 9: Deep learning (optional) (LO3, LO6)

Course and University Policies

Attendance Policy

Students should review the University Class Attendance Policies on the Class Attendance Policies page. The following attendance policy will be applied in this course.

[Insert policy; include description even if course is asynchronous or you do not take attendance]

Generative AI Use in This Course

The use of Generative AI (GenAI) in course assignments and assessments must align with the guidelines established by the instructor. Unauthorized use of GenAI could result in breaches of academic integrity. Instructors are responsible for clearly delineating the permissible uses of GenAI in their courses, underscoring the importance of responsible and ethical application of these tools.

Community Standards within the Office of the Dean of Students articulate the university's stance on academic integrity and scholastic dishonesty. These standards extend to the use of GenAI. Unauthorized or unapproved use of GenAI in academic work falls within the scope of these policies and will be subject to the same disciplinary procedures.

As the instructor for this course, I have adopted the following policy on student use of GenAI.

[Insert policy]

Institutional Policies

UTA students should review the University Catalog and the Syllabus Institutional Policies page for institutional policies and contact the specific office with any questions. The institutional information includes the following policies, among others:

- Drop Policy
- Disability Accommodations
- Academic Integrity
- Electronic Communication

UTA Honor Code

UTA students are expected to adhere to and observe standards of conduct compatible with the University's functions as an educational institution and live by the University of Texas at Arlington's Honor Code. It is the policy of The University of Texas at Arlington to uphold and support standards of personal honesty and integrity for all students consistent with the goals of a community of scholars and students seeking knowledge and responsibility.

Student Support Services

Student Services Page

The Student Services page provides links to many resources available to UTA students, including:

- Academic Success
- Counseling and Psychological Services (CAPS)
- Health Services

- Students with Disabilities
- Veteran Services

Students are also encouraged to check out Career Center resources to enhance their career-readiness, find student employment, search for internships, and more. We encourage Major Exploration and the use of Experiential Major Maps to keep students on track for graduation. Refer to the Graduation Help Desk for more details.

Accessibility of Course Materials

Some course materials, such as PDFs of musical scores, technical drawings, graphs, blueprints, design plans, or artworks (common in fields like drawing, painting, or construction drafting), may not fully comply with all Web Content Accessibility Guidelines (WCAG) requirements.

The University of Texas at Arlington is dedicated to ensuring all students have equal access to information. If you experience any accessibility barriers with course materials, please know that accommodations are available. You can get assistance through the Student Access and Resource (SAR) Center or by contacting your instructor directly. Please don't hesitate to reach out if you need help.

Online Academic Success Guide

Visit the Online Academic Success Guide to explore a list of helpful tips and resources to help you succeed in your online journey.

UTA Health and Wellbeing Resources

UT Arlington is committed to the safety, success, and well-being of our students. To support our community, UTA has established a Community Advocacy, Response, and Engagement (CARE) Team, a dedicated group of campus professionals responsible for helping students who could benefit from academic, emotional, or psychological support, as well as those presenting risks to the health or safety of the community. If you know of someone experiencing challenges, appearing distressed, needing resources, or causing a significant disruption to the UTA community, please submit a CARE Referral by visiting the CARE Team page. You may also submit a referral for yourself if you would like additional support.

UTA students also have access to virtual, on-demand emotional support, appointment-based counseling, advanced psychiatric care, and more. For more information, visit TimelyCare.

NOTE: If a person’s behavior poses an immediate threat to you or someone else, contact UTA Police at 817-272-3003 or dial 911. If you or someone you know needs to speak with a crisis counselor, please reach out to the MAVS TALK 24-hour Crisis Line at 817-272-8255 or the National Suicide and Crisis Lifeline at 988.

Librarian to Contact

Each academic unit has access to Librarians by Academic Subject who can assist students with research projects, tutorials on plagiarism, citation references, as well as support with databases and course reserves.