AMAT 502—Modern Computing for Mathematicians—Spring 2023 Syllabus

Instructor: Prof. Justin M. Curry **Office:** CK 391 **Email:** jmcurry@albany.edu

Course Objectives: *Modern Computing for Mathematicians (AMAT 502)* is designed to take a student with no previous background in computer science or programming to a basic level of competency in programming in Python so that they can carry out the fundamental tasks of a data scientist. Specifically, by the end of the semester the student should be able to: define functions, assign values to variables, work with basic data types such as lists and dictionaries, understand concepts from object oriented programming (OOP) such as inheritance and polymorphism, create a class from as well as subclasses, call methods on objects, control the flow of a program via good programming principles such as abstraction, create scatter plots and plots of mathematical functions using Matplotlib, simulate random phenomena such as flipping a coin, apply machine learning methods such as regression, KNN, K-Means, PCA, logistic regression, SVM, and naive Bayes classifiers. The student will be able to load a data set (provided as a CSV file) as a Pandas data frame and perform basic data cleaning, feature engineering, and apply scikit-learn libraries to analyze said data set. A final project, which consists of a 10-minute recorded presentation is meant to develop presentation skills using some slideshow software, e.g. PowerPoint or Key Note, as well communication skills.

Description of the Course: *Modern Computing for Mathematicians (AMAT 502)* is meant to be the core requirement of the Data Science Masters program here at UAlbany. This course will provide an introduction to programming for students who have never coded before as well as a refresher for more experienced programmers. We will also provide a theoretical and practical introduction into machine learning. At a high-level, this course has three parts:

- 8 lectures on Programming Fundamentals in Python
- 4 lectures on Numpy, SciPy and Statistics
- 11 lectures on the machine learning using Pandas and Scikit-Learn as well as its conceptual underpinnings.

The remaining time is for group projects and presentations. A lecture outline is on Page 3.

Pre-Requisites: Basic undergraduate mathematics such as calculus, linear algebra, probability and stats.

Lecture Place and Times: This is an asynchronous online course. You will need to download the Global Protect VPN (first link below) to access the appropriate server for your class when you're off campus. Your login is your netID **all lower case** and your usual password.

- https://wiki.albany.edu/pages/releaseview.action?pageId=77139467
- https://amat502-8784-fa23.its.albany.edu/
- https://amat502-8785-fa23.its.albany.edu/
- Backup Option: https://github.com/jmc42/AMAT502
- We also have a Discord Server. See brightspace.albany.edu for the invite link.

Office Hours: MW 11:40-1pm ONLINE

- https://albany.zoom.us/j/98123214907?pwd=eURYbHVyQlhVWnF0anMyR21HNmJHZz09
- Email Brian Heller bheller@albany.edu if you have more questions.

Grading Schema:

- 40% Determined by 6 problem sets, and 1 programming project; denoted ♣.
- 30% Determined by a take-home 72-hour midterm; denoted by ♢.
- 30% Determined by an individual final project and proposal; denoted by ♠.

Academic Integrity: Please familiarize yourself with UAlbany's *Standards of Academic Integrity* page: https://www.albany.edu/studentconduct/standards_of_academic_integrity.php

I take violations of academic integrity very seriously. Programming is a subject where it is too easy to cheat yourself of valuable learning opportunities. It is permissible to consult existing questions and answers on *Stack Overflow* or *Geeks for Geeks* or similar programming help websites, but you CANNOT post one of the questions from the HW or projects as a new question to one of these sites or use Chegg or a similar "homework help" service. You are, however, permitted to ask your classmates or me or the TAs for help, but you should make sure you've tried to figure it out on your own for 10-15 minutes at least.

N.B. The single most important thing to NOT do is to copy large blocks of code from a classmate or the internet. If you are caught copying and pasting someone else's code the following penalty system will apply:

(1) For the first offense, **a zero for the question that you copied on.**
(2) For the second offense, **a zero for the assignment and I file a Violation of Academic Integrity Report (VAIR).**
(3) For the third or later offense, **you get a letter grade reduction (or possibly fail) and I refer you to community standards, which could result in expulsion from UAlbany.**

Course Materials: Jupyter Notebooks cover all the necessary material for this class as well as videos for each lecture. Please watch these ahead of time.

- GitHub Back-Up of Lectures: `https://github.com/jmc42/AMAT502/`
- YouTube Lectures: `https://www.youtube.com/channel/UC3HHUGPjUfyH0YmkYXfFMbQ`
- Required Text: *Introduction to Computation and Programming Using Python* by Guttag, 3rd ed.
- Additional Free Text: `https://github.com/jakevdp/PythonDataScienceHandbook`

Attendance and Missed Work Policy:

This is an asynchronous online course, so there is no required attendance. However, I will run regular office hours, which you may benefit from attending. Although this course is asynchronous, deadlines for homework and projects are strict. This is to make sure you keep up with the material and so that I can provide you with feedback. ***I will take off 2 points off for every day an assignment is late, unless I grant you an extension, which must be negotiated before the due date.***

Lecture Outline:

The lectures are pre-recorded and available on YouTube. Search for "AMAT502". The following is the link to the channel. Please Like and Subscribe!

https://www.youtube.com/channel/UC3HHUGPjUfyH0YmkYXfFMbQ

- Lec 1 = Introduction to Data Science + Basic Numerical Operations,
- Lec 2 = Conditionals and Loops,
- Lec 3 = Bisection Search and Functions, Quick Discussion of Big-O Notation
- Lec 4 = Functions and Recursion,
- Lec 5 = More Recursion, Strings and Lists
- Lec 6 = Data Types and Edit Distances
- Lec 7 = "One Liners" and Intro to Object Oriented Programming
- Lec 8 = More OOP and Data Structures
- Lec 9 = OOP for Mortgages and Matplotlib
- Lec 10 = NumPy Array Operations + Intro to Randomness
- Lec 11 = Basic Probability Distributions Review
- Lec 12 = Hypothesis Testing and the CLT
- Lec 13 = Into to Machine Learning: Regression
- Lec 14 = Intro to ML: Classification and Clustering Overview,
- Lec 15 = Catch-Up Lecture + Intro to Pandas
- Lec 16 = K-Means Clustering
- Lec 17 = Principal Component Analysis
- Lec 18 = MNIST via K-means and K-Nearest Neighbors
- Lec 19 = Model Validation and Bias-Variance Tradeoff
- Lec 20 = Naive Bayes Classifier
- Lec 21 = Maximum Likelihood and Logistic Regression
- Lec 22 = SVM: From Hyperplanes to Kernels
- Lec 23 = Decision Trees and Entropy
- Lec 24 = Data Science Reconsidered + Intro to TDA

Look above for the approximate content of the lectures listed below.

Monday	Wednesday
Aug 21st Lecture 1	**Aug 23rd** Lecture 2
Aug 28th Lecture 3 **♣PS00 Due Monday 8/28♣**	**Aug 30th** Lecture 4
Sep 4th ♡*LABOR DAY = NO CLASS*♡	**Sep 6th** Lecture 5 **♣PS01 Due Friday 9/8♣**
Sep 11th Lecture 6	**Sep 13th** Lecture 7 **♣PS02 Due Friday 9/15♣**
Sep 18th Lecture 8	**Sep 20th** Lecture 9
Sep 25th Lecture 10	**Sep 27th** Lecture 11 **♣PS03 Due Friday 9/29♣**
Oct 2nd Lecture 12	**Oct 4th** **♢Distribute Practice Midterm♢**
Oct 9th ♡*FALL BREAK*♡	**Oct 11th** **♢MIDTERM♢**
Oct 16th Lecture 13	**Oct 18th** Lecture 14
Oct 23rd Lecture 15	**Oct 25th** Lecture 16 **♣PS04 Due Friday 10/27♣**
Oct 30th Lecture 17	**Nov 1st** Lecture 18 **♣PS05 Due Friday 11/3♣**
Nov 6th Lecture 19	**Nov 8th** Lecture 20
Nov 13th Lecture 21	**Nov 15th** Lecture 22 **♣Project K-Means Due Friday 11/17♣**
Nov 20th Lecture 23	**Nov 22nd** ♡*THANKSGIVING*♡
Nov 27th **♠Project Proposals Due Monday 11/27♠**	**Nov 29th** ♠*Final Projects—Due Monday 12/11!*♠
Dec 4th ♠*Final Projects—Due Monday 12/11!*♠	**Dec 6th** ♡*NO MORE CLASSES*♡

FINAL PROJECTS ARE DUE MONDAY DECEMBER 11th!