MAT/PHY 231 Think Like a Data Scientist Syllabus
Agnes Scott College, Spring 2023, MW 12:15–1:30 in BSC 102W
Jim Wiseman, jwiseman@agnesscott.edu, Buttrick 331

Office hours: Mon 1:45-2:45, Tues 2:00-3:00, and by appointment.

Textbook: We're using a free online textbook called How to Think like a Data Scientist. It is available at https://runestone.academy/ . You must go to the website and register for the class; use our class name, ASC231Spring2023, when you register. Always use this account for this class!
We will also use the book Weapons of Math Destruction, by Cathy O'Neil.

Course Philosophy
This course introduces students to the importance of gathering, cleaning, normalizing, visualizing and analyzing data to drive informed decision-making, no matter the field of study. Students will learn to use a combination of tools and techniques, including spreadsheets, SQL and Python to work on real-world datasets using a combination of procedural and basic machine learning algorithms. They will also learn to ask good, exploratory questions and develop metrics to come up with a well-thought-out analysis. Presenting and discussing an analysis of datasets chosen by the students will be an important part of the course. Like PHY/MAT 131, this course will be "flipped," with content learned outside of class and classroom time focused on hands-on, collaborative projects.

Goals for the Course
Students in this course will learn how to:
1. Locate, compare and select online sources of data for analysis
2. Create algorithms and programs that clean, normalize and visualize data
3. Plan basic machine learning algorithms
4. Assess real world datasets using spreadsheets, SQL, and Python

Homework and Class Preparation
In this class you will have daily assignments, including reading assignments from the online textbook. The online textbook will provide a flipped classroom style of learning, meaning that I will not be spending a lot of time lecturing; your primary source for new information will be the book. The reading assignments are active assignments; you cannot just passively read the material you must interact with it. I will be able to gauge your level of interaction with the material from the online system.

Tips for success

- Always always feel free to ask questions!
- The materials in this class all build on each other. Therefore it is really important to keep up. You can't ignore one topic and hope it will go away as once we have introduced a new idea we will continue to use it throughout the course. If you do get behind let's talk so we can devise a strategy to get you caught up.
- It is almost inevitable that technical problems will arise. When they do, don't panic. I will not punish someone with a bad grade due to a technical glitch.

- In Class – We will be using computers in class every day, so be wary of distractions. It is very tempting to check Facebook, email, twitter, or whatever when the computer is in front of you. Don't do it!! It not only distracts you but everyone around you.

- Grading
  Your grade in the course will be determined using the following percentages. Participation will include attendance and activity/participation in class.

    Reading Assignments (including inline questions and discussion) 25%
    Class work and home work 25%
    Midterm Project 25%
    Final project 25%
    Total 100%

Autograded Work

Much of the homework will be automatically graded, you will get immediate feedback as you run and test your work. You are encouraged to keep working and to get as many of the automatic tests to pass as you can. You WILL NOT be penalized for making lots of tries.

Software
Early in the semester, we will be using Google Sheets for our data analysis. Later in the semester, we will use Jupyter notebooks on Google Collab.

Honor code and group work
All students are expected to follow the honor code throughout the semester; all exams and assignments should be pledged.
I strongly encourage you to work on the homework in groups. I suggest that you work on the problems by yourself first, making a note of anything giving you trouble; then meet with your group and work through the remaining problems together; and finally submit the solutions by yourself. Every group member must submit her own solutions independently; just copying the group's answers is plagiarism and is unacceptable.
The Agnes Scott College honor code embodies an ideal of character, conduct, and citizenship, and is an important part of the College's mission and core identity. This applies especially to academic honesty and integrity. Passing off someone else's work as your own represents intellectual fraud and theft, and violates the core values of our academic community. To be honorable, you should understand not only what counts as academic dishonesty, but also how to avoid engaging in these practices. You should:
• review each course syllabus for the professor's expectations regarding course work and class attendance.
• attribute all ideas taken from other sources; this shows respect for other scholars. Plagiarism can include portraying another's work or ideas as your own, buying a paper online and turning it in as if it were your own work, or not citing or improperly citing references on a reference page or within the text of a paper.
• not falsify or create data and resources or alter a graded work without the prior consent of your professor. This includes making up a reference for a works cited page or making up statistics or facts for academic work.
• not allow another party to do your work/exam, or submit the same or similar work in more than one course without permission from the course instructors. Cheating also includes taking an exam for another person, looking on another person's exam for answers, using exams from previous classes without permission, or bringing and using unauthorized notes or resources (i.e., electronic, written, or otherwise) during an exam.

• not facilitate cheating, which can happen when you help another student complete a take home exam, give answers to an exam, talk about an exam with a student who has not taken it, or collaborate with others on work that is supposed to be completed independently.
• be truthful about the submission of work, which includes the time of submission and the place of submission (e.g., e-mail, online, in a mailbox, to an office, etc.).
Because of the centrality of the Honor Code to our campus life, penalties result from dishonest conduct. In academic courses, these penalties can range from failure of the assignment to expulsion from the college. You should speak with your professors if you need clarification about any of these policies.

Late work: Late work won't be accepted. Everything that we do in this class builds on what we've done previously, so it is extremely important that you keep up with deadlines.

Attendance and participation: I expect you to be at every class meeting on time, unless you've talked to me about having to be absent for technological or other reasons. However, tardiness or absence will have no (direct) effect on your grade, unless you miss a day of presentations.

Course evaluation: Your feedback on the course is extremely valuable to me, the math department, and the administration. In particular, I take your comments very seriously and use them to improve the course the next time I teach it. You are responsible for completing an evaluation of the course at the end of the semester.

Title IX: Agnes Scott is here to help you if you have experienced any form of sexual harassment or violence, dating or domestic violence, or stalking. Please talk to any faculty or staff member with whom you feel comfortable. Faculty and staff members want to support you and have been trained to help. They will also inform the Title IX office so that you learn about options available to you. If you do not want college administrators to know what you have experienced, you may talk to the chaplain, as well as nurses or counselors in the Wellness Center with complete confidentiality. They will not tell anyone what you share with them unless you give your express permission. You may contact the Title IX Coordinator directly at T9Coordinator@agnesscott.edu.

Inclusion: This course adheres to the principles of diversity and inclusion integral to the Agnes Scott community. We respect people from all backgrounds and affirm people's decisions about gender expression and identity. Please let me know your preferred name or gender pronoun if different from the class roster.

ADA: Agnes Scott College seeks to provide equal access to its programs, services and activities for people with various abilities. If you will need accommodations in this class, please contact the Office of Academic Advising and Accessible Education (404-471-6150) to complete the registration process. Once registered, please contact me so we can discuss the specific accommodations needed for this course.

Schedule
The following course schedule is subject to change! See Canvas for more details.

Date	In Class	Reading and Practice (complete before class)
Wed 1/11	Intro, Collecting Class Data 1	none
Mon 1/16	**MLK Day - no classes**	
Wed 1/18	Collecting Class Data 2	Read Chapter 1
Mon 1/23	Happiness Data 1	Read Chapter 2 (2.1-2.2)
Wed 1/25	Happiness Data 2	Read Chapter 2 (2.3)
Mon 1/30	Happiness Data 3	Read Chapter 2 (2.4)
Wed 2/1	WMD Discussion	Read WMD (1-3) Kaggle/Intro
Mon 2/6	Python Review	Read Chapter 4 (4.1-4.2)
Wed 2/8	Jupyter Notebooks, start Pandas	Read Chapter 4 (4.3 and 4.5) *Finalize choice of midterm dataset*
Mon 2/13	Pandas 1	Read Chapter 5 (5.1-5.4)
Wed 2/15	Pandas 2	Read Chapter 5 (5.5-5.6)
Mon 2/20	Working with Midterm Dataset	Read Chapter 6
Wed 2/22	Data Ethics	Read WMD (4-6)
Mon 2/27	Midterm Presentations 1	
Wed 3/1	Midterm Presentations 2	
3/6-3/17	**Spring break/Journeys/Peak Week**	
Mon 3/20	WMD/Data Ethics Discussion	Read Chapter 7 Read WMD (7-8)
Wed 3/22	Textual analysis 1	Read Chapter 8 (8.1-8.2)
Mon 3/27	Textual analysis 2	Read Chapter 8 (8.3)
Wed 3/29	Textual analysis 3	Read Chapter 8 (8.4-8.5)
Mon 4/3	Textual analysis 4	Review Chapter 8 (8.5)
Wed 4/5	Textual analysis 5	Read Chapter 8 (8.6)
Mon 4/10	Textual analysis 6	Read Chapter 8 (8.7)
Wed 4/12	Predictive analytics 1	Read Chapter 9 (9.1-9.8)

Date	In Class	Reading and Practice (complete before class)
Mon 4/17	Predictive analytics 2	Read Chapter 9 (9.9-9.11)
Wed 4/19	Project work/WMD Final Discussion	Finish WMD
Mon 4/24	Regression, Scikit Learn	
Wed 4/26	Project work	
Mon 5/1	Final Presentations 1	
Wed 5/3	Final Presentations 2	