

DATA 601: Introduction to Data Science

Class

- Wed: 7:10 - 9:40 PM
- Online: (Link will be shared via Blackboard)

Instructor Mehmet Sarica

Office Hours: <https://calendly.com/msarica1/15min>

E-mail: mehmet.sarica@umbc.edu

Q&A and Discussions

We will be using Piazza for email, questions, and class discussion. Piazza is highly catered to getting you help fast and efficiently from classmates and myself. Rather than emailing questions, I highly encourage you to post your questions on Piazza. Find our class page by following the link: piazza.com/umbc/spring2021/ds601

In addition, I will be available after class for questions and help. I check my email daily and mostly in the evenings. If your question is something about your code and if it requires some context please schedule an appointment.

Course Description

The goal of this class is to give students an introduction to and hands on experience with all phases of the data science process using real data and modern tools. Topics that will be covered include data formats, loading, cleaning, data storage in relational & non-relational stores and data visualization.

Learning Objectives

Upon completion, students will be able to:

- utilize Python data structures such as lists, dictionaries, strings.
- use common Python modules such as random, pandas and matplotlib.
- understand issues relating to acquisition, cleaning and loading of data.

- understand the basics of how data can be presented and visualized.
- use common data formats such as CSV, JSON, XML.
- load & clean data
- handle missing data and outliers.
- analyze data & present meaningful visualization.
- develop a solution for a problem.

Covid-19

Covid-19 Due to the global pandemic, our classes will be held remotely this semester. Please

see <https://docs.google.com/document/d/1xWWGAR8qEzKYr7qaVHoEhvO6lyXlyn6M3M7EFZPJQgA/> for UMBC Policies and Resources during COVID-19.

Github Repository

I will use https://github.com/msaricaumbc/DS601_Spring21 as our course webpage. You can check this page for updates and important information related to the course.

Textbooks

- A Whirlwind Tour of Python by Jake VanderPlas <https://jakevdp.github.io/WhirlwindTourOfPython/>
- Python Data Science Handbook by Jake VanderPlas <https://jakevdp.github.io/PythonDataScienceHandbook/>

Course Format and Assignments

Students will complete assigned homework assignments, readings, projects. This course incorporates a variety of hands-on labs and practical exercises to engage students and prepare them for challenges they may encounter in a workplace.

The final project will provide students opportunity to showcase what they have learned in a format similar to what they will encounter in a professional work setting.

Grading Criteria

Students are expected to participate in class discussions. Student participation in online classes is encouraged & expected; watching recorded lectures outside of class does not count towards attendance!

Homework assignment will be mostly due Sunday and it will have +5 point incentive if turned in by Friday. For homework due dates please check the weekly schedule below.

Course Work

Course work	Grade Distribution
Attendance & Participation	10%
Homework	30%
Midterm Project	30%
Final Project	30%

Grades

Percentile	Letter Grade
93-100	A
90 - 92	A-
87-89	B+
83-86	B
80-82	B-
77-79	C+

Percentile	Letter Grade
------------	--------------

73-76	C
-------	---

70-72	C-
-------	----

60-69	D+
-------	----

0- 59	F
-------	---

Weekly Schedule

The following is a tentative schedule. We may update it as we go.

Wk	Date	Topic	Readings	Presentations	Homework
1	Jan 27	Welcome, Logistics	Overview Data Science Environment Setup Google Colab Tutorial		TBD
2	Feb 3	Python Basics			TBD
3	Feb 10	Data Formats			TBD
4	Feb 17	Acquiring and Storing Data			TBD
5	Feb 24	Data Cleanup part-I			TBD

Wk	Date	Topic	Readings	Presentations	Homework
6	Mar 3	Data Cleanup Part-II			TBD
7	Mar 10	Data Exploration and Analysis			TBD
8	Mar 17	Spring Break	No Class		TBD
9	Mar 24	Data Exploration and Analysis Part-II			TBD
10	Mar 31	Presenting Your Data			TBD
11	Apr 7	Web scraping			TBD
12	Apr 14	API's			TBD
13	Apr 21	Automation and Scaling			TBD
14	Apr 28				TBD
15	May 5				TBD

Wk	Date	Topic	Readings	Presentations	Homework
16	May 12	Final Projects			Group Presentations

Course Policies

Assignments

UMBC provides a range of writing assistance, which can be found in the following:

- The Writing Center: <http://lrc.umbc.edu/tutor/writing-center/>
- Research Guides & Tutorials: <http://lib.guides.umbc.edu/tutorial>

Failure to follow guidelines for each assignment, including the required format, style, length, and submission may result in at least one-letter-grade reduction on the assignment depending on the type or number of transgressions.

Late Assignments

NO late/incomplete assignments will be accepted! Emergency situations will be handled on a case by case basis with appropriate justification and/or documentation.

Incomplete grades are granted only for extenuating circumstances and your request is made before the last week of class.

Academic Integrity

By enrolling in this course, each student assumes the responsibilities of an active participant in UMBC's scholarly community in which everyone's academic work and behavior are held to the highest standards of honesty. Cheating, fabrication, plagiarism, and helping other to commit these acts are all forms of academic dishonesty, and they are wrong. Academic misconduct could result in disciplinary action that may include, but is not limited to failure, suspension or dismissal.

Refer to the UMBC policy on Academic

Integrity: <http://catalog.umbc.edu/content.php?catoid=17&navoid=879#academic-integrity>.

Student Disability Services (SDS)

UMBC is committed to eliminating discriminatory obstacles that may disadvantage students based on disability. Services for students with disabilities are provided for all students qualified under the Americans with Disabilities Act (ADA) of 1990, the ADAAA of 2009, and Section 504 of the Rehabilitation Act who request and are eligible for accommodations. The Office of Student Disability Services (SDS) is the UMBC department designated to coordinate reasonable accommodations that would allow students to have equal access and inclusion in all courses, programs, and activities of the University. If you have a documented disability and need to request academic accommodations, please register with the Office of Student Disability Services (SDS) as soon as possible. To begin the registration process please visit the SDS website and review the registration information, including disability documentation guidelines and how to submit the SDS registration form online using the confidential data management software called Accommodate <https://sds.umbc.edu/accommodations/registering-with-sds/>. Once accommodations have been approved, you and your instructors will be notified via an emailed accommodation letter from the SDS office. Both the SDS office and Shady Grove's Center for Academic Success(CAS) will work with you to ensure you receive the approved accommodations. If you have any questions or concerns, please contact the Office of Student Disability Services via disAbility@umbc.edu or phone at 410-455-2459. Please note that accommodations are not retroactive and begin once SDS sends an approved accommodation letter. For more information on the services CAS provides, please contact Mary Gallagher (maryg@umd.edu) or visit <https://shadygrove.umd.edu/student-services/center-for-academic-success>.