

Data 601: Introduction to Data Science

Fall 2020

General Information

Meeting Times and Location

Thursday, 4:30pm-7:00pm, Online

Instructor

Murat Guner

Email

mguner@umbc.edu

Office Location & Hours

time TBD

I check my email daily and generally will respond to questions in the evening hours. I am available before class for questions and help. Outside of that timeframe, please email me to schedule an appointment.

Covid-19

Due to the global pandemic, our classes will be held remotely this semester. Please see [this Google doc](#) for UMBC Policies and Resources during COVID-19.

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, and cleaning; data storage in relational and non-relational stores; data analysis using supervised and unsupervised learning using Python; data visualization.

Prerequisite

Students must be enrolled in the Data Science Program. Other students may be admitted with instructor permission.

Course Learning Objectives

Upon completion, students will:

- Understand issues relating to acquisition, cleaning and loading of data.
- Be able to perform data analysis using Python.
- Understand the basics of how data can be presented and visualized.
- Understand issues involved when the analysis scales up to Big Data.
- Understand pros and cons of the different data analysis methods.
- Understand data science projects life cycle.
- Be aware of ethical and legal aspects of a data science project.

Course Materials

Optional Texts

- "Think Like a Data Scientist: Tackle the data science process step-by-step" by Brian Godsey. Manning Publications. You can download the book from this [link](#).
- Data Science for Business: What You Need to Know About Data Mining and Data-Analytics Thinking by Foster Provost and Tom Fawcett, O'Reilly Media. Here is a [link](#) for the book.
- "Python Data Science Handbook" by Jake VanderPlas. O'Reilly Media. You can access the book from this [link](#).
- Practical Data Science with Python 3: Synthesizing Actionable Insights from Data by Ervin Varga. Apress. Here is a [link](#) for the book.
- "Data Wrangling with Python: Tips and Tools to Make Your Life Easier" by Jacqueline Kazil and Katharine Jarmul. O'Reilly Media

Please review options at the [UMBC library](#) and PDFs available online.

Recommended Software and Hardware

- Web browser capable of running Jupyter Notebooks
- A computer with sufficient internet speed for online lectures. Make sure that your computer has video and microphone access.

Course Format and Assignments

The students will complete assigned homework, a midterm project and a final project with presentation. This course incorporates a variety of hands-on labs and practical exercises to engage students and prepare them for challenges they may encounter in the 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.

Course Communication

I use the Slack channel at <https://umbcdatasci.slack.com/messages> for addressing questions.

Grading Criteria

Students are expected to participate in class discussions.

Course work

Attendance and Participation
Homework
Midterm project
Final Project and Presentation

Grade Distribution

10%
30%
30%
30%

Final Grade will be computed as follows:

90-100%	A
80 to 89%	B
70 to 79%	C
60 to 69%	D
<60	F

Equity and Inclusion Policies

Diversity Statement

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength and benefit. It is my intent to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

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>.

Title IX Statement:

Any student who has experienced sexual harassment or assault, relationship violence, and/or stalking is encouraged to seek support and resources. There are a number of resources available to you. Please see [this website](#) for recently updated UMBC Policies and Resources during COVID-19.

With that said, as an instructor, I am considered a Responsible Employee, as per [UMBC's Interim Policy on Prohibited Sexual Misconduct, Interpersonal Violence, and Other Related Misconduct](#). This means

that while I am here to listen and support you, I am required to report disclosures of sexual assault, domestic violence, relationship violence, stalking, and/or gender-based harassment to the University's Title IX Coordinator. The purpose of these requirements is for the University to inform you of options, supports, and resources.

You can utilize support and resources even if you do not want to take any further action. You will not be forced to file a police report, but please be aware, depending on the nature of the offense, the University may take action.

If you need to speak with someone in confidence about an incident, UMBC has the following Confidential Resources available to support you:

- [The Counseling Center](#): [410-455-2742](#) (M-F 8:30 a.m. – 5 p.m.)
- [University Health Services](#): [410-455-2542](#) (M-F 8:30 a.m. – 5 p.m.)
- For after-hours emergency consultation, call the [police](#) at [410-455-5555](#)

Other on-campus supports and resources:

- [The Women's Center](#) (available to students of all genders): [410-455-2714](#) (M-Th 9:30 a.m. – 6 p.m., F 9:30 a.m. – 4 p.m.)
- [Title IX Coordinator](#): [410-455-1606](#) (9 a.m. – 5 p.m.)

Child Abuse and Neglect

Please note that Maryland law requires that I report all disclosures or suspicions of child abuse or neglect to the Department of Social Service and/or the police.

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

Late/incomplete assignments will be accepted if an extension has been agreed to in advance. Emergency situations will be handled on a case by case basis with appropriate justification 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>.

Course Outline

Subject to revision

Week 1	Introduction to course and data science
Week 2	Python Basics
Week 3	Data Visualization/Cleaning
Week 4	APIs and Web scraping
Week 5	Data Science Life Cycle
Week 6	Supervised Learning Problems - Regression
Week 7	Classification Problems
Week 8	Evaluation/ Model Selection
Week 9	Unsupervised Learning problems - Clustering
Week 10	Unsupervised Learning problems
Week 11	Working with Time Series
Week 12	Data Ethics and Legality
Week 13	Misc.
Week 14	No class (Thanksgiving)
Week 15	Review and Closing Remarks

