CICS 108 - Foundations of Data Science

Instructor
Purity Mugambi pmugambi@umass.edu
(Email is the best way to reach me. Allow 24 hours for a response.)

Course Description
The field of Data Science encompasses methods, processes, and systems that enable the extraction of useful knowledge from data. Foundations of Data Science introduces core data science concepts including computational and inferential thinking, along with core data science skills including computer programming and statistical methods. This course presents these topics in the context of hands-on analysis of real-world data sets, including economic data, document collections, geographical data, and social networks. The course also explores social issues surrounding data analysis such as privacy and design.

Assignments and Grading
All major assignments and the extra credit will be submitted via [Gradescope](#) and all feedback and grades will be communicated via [Moodle](#). I will cover the submission process in class. Feedback will not be posted prior to the due date, but students may submit an assignment as early as they wish.

Your success in this course is important. Your instructor has provided several resources and opportunities for you to learn the contents of this course. Some of those include: numerous in-class demonstrations and hands-on activities, lab sessions where you can further practice the content taught during lectures, instructor and teaching assistant (TA) office hours (OHs) where you can seek more help and guidance. These resources are in addition to many other resources student success provided by the college and the University, such as your academic advisor. You are encouraged to take advantage of these resources to ensure you succeed in this course.

Throughout the semester, your instructor will communicate with Student Success & academic advisors regarding your progress in the course. If you are contacted, please consider scheduling appointments such as tutoring or academic advising and talk with your professor. Referrals are not punitive and are meant to assist you in connecting with resources at UMass. Please email academicalert@umass.edu if you have any questions or need assistance connecting with resources.

- **Homework = 35%** There are 10 assignments in this course, one for each Module. The assignment will be **released every Tue after the lecture**, and will be **due next Tue before the lecture**. Grades of the assignments will be available on Gradescope **within 2 weeks after the assignment due date**. Regrade requests will be accepted no later than **one week** after grades are released on Gradescope. After that, all assignment grades will be posted on Moodle and deemed final. **LATE homeworks will NOT be graded.**

- **Labs = 20%** There are 10 labs in this course, one for each Module. Labs are a one hour session where students engage in hands-on coding to answer questions on the material taught that week during class lectures. Students will be grouped in 2s/3s each week to work on the lab exercise together. When you submit your lab on gradescope, make sure to add the names of your groupmates. **Attendance is important** for labs because **all work will be completed during the session**. If you cannot make it to a lab session let your instructor know in advance (at least a day in advance) so that they can make other arrangements for you. **LATE labs will NOT be graded.**
- **Midterm exam = 20%** The midterm exam will be on **Oct 13 8:30am - 9:45am, i.e. during lecture time**. If more time is needed to complete the exam, **Oct 14 9:05 - 9:50am, i.e., lab time will** be used to complete it. Students have to be in the lab room in person to continue the exam. **This will be an open book exa**m. The questions will mostly ask the students to write code.
- **Final exam = 20%** The final exam **will be take-home, and open book**. It will be released on **Dec 6 after class,** and will be due on **Dec 14 11:59pm EST**. However, once you start the exam, **you will have 4 hours** to complete and upload it to Gradescope.

Attendance – Attendance is highly encouraged for this class because most of the material will be learned through hands-on activities. This is especially important for Labs where **all work** will be done within the lab hour.

Announcements – Moodle and email will be used for conveying important class announcements.

Accommodation Policy – Students with a documented disability are encouraged to communicate their needs to the instructor at their earliest convenience. If you aren't sure if you qualify for an accommodation, or if you have any questions about your right to accommodations, please contact the office of Disability Services: http://www.umass.edu/disability/students

Cell Phones and other digital devices not used for class discussion/– To avoid distracting the class, please put the devices you are not using away or keep them in silent mode. I expect you to use all your devices responsibly so as not to distract yourself or the class from learning.

Inclusion Policy – In this course, each voice in the classroom has something of value to contribute. Please take care to respect the different experiences, beliefs and values expressed by students and staff involved in this course. We support UMass Amherst's commitment to diversity, and welcome individuals of all ages, backgrounds, citizenships, disability, sex, education, ethnicities, family statuses, genders, gender identities, geographical locations, languages, military experience, political views, races, religions, sexual orientations, socioeconomic statuses, and work experiences.

I reserve the right to make changes to this syllabus as needed to incorporate current events and to better support positive student outcomes.

Week	Date	Activity	Module
1	9/6 Tu	Lecture 1: Introductions, course policies & syllabus, Cause & Effect HW1 assigned	1
1	9/8 Th	Lecture 2: Association vs Causation	1
1	9/9 F	Lab 1	1
2	9/13 Tu	**HW1 due at 8am** Lecture 3: Data types - Numbers, Strings, Comparisons & Booleans, Ranges HW2 assigned	2
2	9/15 Th	Lecture 4: Data types - Arrays, Lists, Tables	2
2	9/16 F	Lab 2	2
3	9/20 Tu	**HW2 due at 8am** Lecture 5: Data visualization HW3 assigned	3
3	9/22 Th	Lecture 6: Data visualization	3
3	9/23 F	Lab 3	3
4	9/27 Tu	**HW3 due at 8am** Lecture 7: Functions HW4 assigned	4
4	9/29 Th	Lecture 8: Pivot tables	4
4	9/30 F	Lab 4	4
5	10/4 Tu	**HW4 due at 8am** Lecture 9: Randomness HW5 assigned	5
5	10/6 Th	Lecture 10: Simulation	5
5	10/7 F	Lab 5a	5
6	10/11 Tu	Lecture 11: Sampling	5
6	10/13 Th	Lecture 12: Distributions	5
6	10/14 F	Lab 5b	5
7	10/18 Tu	**HW 5 due at 8am** Midterm Review	N/A

7	10/20 Th	**NO LECTURE - Midterm exam** (in class, during usual lecture time)	N/A
7	10/21 F	**NO LAB, extra time reserved for midterm exam**	N/A
8	10/25 Tu	Lecture 13: Testing hypotheses HW6 assigned	6
8	10/27 Th	Lecture 14: Comparing 2 samples	6
8	10/28 F	Lab 6	6
9	11/1 Tu	**HW6 due at 8am** Lecture 15: Estimation 1 HW7 assigned	7
9	11/3 Th	Lecture 16: Estimation 2	7
9	11/4 F	Lab 7	7
10	11/8 Tu	**HW7 due at 8am** Lecture 17: Why the mean matters 1 HW8 assigned	8
10	11/10 Th	Lecture 18: Why the mean matters 2	8
10	11/11 F	**NO LAB - VETERANS DAY**	N/A
11	11/15 Tu	**HW8 due at 8am** Lecture 19: Prediction HW9 assigned	9
11	11/17 Th	Lecture 20: Inference for regression	9
11	11/18 F	Lab 8	8
12	11/22 Tu	Friday schedule followed - Lab 9	9
12	11/24 Th	**NO CLASS - THANKSGIVING BREAK**	N/A
12	11/25 F	**NO LAB - THANKSGIVING BREAK**	N/A
13	11/29 Tu	**HW9 due at 8am** Lecture 21: Classification HW10 assigned	10
13	12/1 Th	Lecture 22: Classification	10
13	12/2 F	Lab 10a	10
14	12/6 Tu	**HW10 due at 8am**	10

		Lecture 23: Updating predictions **Final exam released**	
14	12/8 Th	Lecture 24: Recap, and final exam review	10
14	12/9 F	Lab 10b	10
15	12/14 W	**Final exam due 11:59pm**	**N/A**