

GMS 6803: Data Science for Clinical Research (3 credit hours)

Fall 2024

LOCATION: HPNP 1101

CLASS HOURS: Tuesday: 12:50 – 3:50 pm

INSTRUCTORS:

Mamoun Mardini, MS, PhD

Assistant Professor

Malachowsky Hall for Data Science and Information Technology 7121

Email: malmardini@ufl.edu

Protiva Rahman, PhD

Research Assistant Professor

Malachowsky Hall for Data Science and Information Technology 7120

Protiva.rahman@ufl.edu

COURSE OVERVIEW:

GMS 6803 provides students with an introduction to a wide range of concepts and techniques in data science as they apply to biomedical and clinical research. Data is an essential component of biomedical and clinical research. It is critical for students to understand and gain practical experience with the entire life cycle of data, from data collection to data analysis to the dissemination and archiving of valuable results. In this course, students are introduced to the broad landscape of data science for biomedical and clinical research: learn how to design and implement computerized databases for data collection, perform basic query and reporting operations, prepare databases for analytical tasks, perform quality assurance procedures, and understand basic data analytical methods and approaches. Further, today's researchers are mining big datasets for patterns and trends that lead to new hypotheses and new discoveries. This course also aims to give students insight into tools, methods, and approaches for big data analytics in the biomedical domain. This course will be a foundation for students who are interested in becoming the next generation of biomedical data scientists.

COURSE OBJECTIVES:

Teaching methods include lecture, discussion, and hands-on data assessment, analysis, and presentation. The goals of the course are:

- To provide basic understanding of the steps in the life cycle of data in biomedical and clinical research.
- To familiarize students with basic principles of data management.
- To deepen students' understanding of data structure, data standards, and data quality issues in data-intensive biomedical researches.
- To help students understand ethic and legal issues when dealing with biomedical and clinical data.
- To introduce the concepts of big data, and the associated tools, methods and approaches.
- To give students access to exploratory data analysis techniques and tools.

TEXTBOOKS/READING MATERIALS:

The instructors will distribute lecture handouts when necessary.

OFFICE HOURS:

Office hours are by request. Please email the instructors for an appointment in advance. We can likely address the questions over email. If not, please make an appointment, and we will try to accommodate your schedule.

PREREQUISITES:

Although the main goal of the course is not to teach programming, a fair amount of coding is required to complete the coursework. Students are required to have basic computer-related skills and knowledge (e.g., operating system(s)). Prior experience with one or more data management software systems (e.g., MS Excel, Access, SQL, etc.) is required. Prior experience with one or more programming languages (e.g., general programming languages such as Java and C/C++; scripting languages such as Python and Ruby; and statistical computing languages such as SAS, SPSS, R, and Matlab) is required.

GRADE COMPOSITION:

Attendance: 5%

Homework assignments & quizzes: 40%

Midterm: 25%

Final (project report): 30%

Homework assignments:

Assignments consist of small programming projects, discussions, and writing critiques for research papers. Students are expected to finish one to two small programming projects. These programming exercises are often simple, and their implementations are easy to find online. You can use these online implementations as references. However, you will be penalized if you merely copy others' work. We reserve the right to ask you to explain your code line-by-line.

Assignment rules: You are required to comply with these rules.

- Your assignment must be turned in no later than 10 am on the day of the class.
- Late homework assignments will be penalized by taking 1 point off per late day.
- No handwritten assignment. All assignments must be submitted electronically through the online system – no emails.
- DO NOT COPY OTHERS' HOMEWORK. There is zero tolerance, and both will receive 0.
- Searching for a solution on the web—and then submitting it as your answer for a homework assignment—will be considered a violation.

Course project:

The final product of the course is a course project, which consists of 30% of the overall grade. Each student is required to complete a course project. You can collaborate with other students as a team. However, each team can have up to four (4) members. Exceptions can only be made with a written explanation and are subject to the instructors' approval. Please clearly delineate the roles and responsibilities of each team member. Your final grade for the course project will be adjusted based on your contribution.

Midterm:

This will be a closed-book examination to assess your understanding of the main concepts in the class.

Attendance policy:

Class attendance is mandatory. Excused absences follow the criteria of the UF Graduate Catalogue (e.g., illness, serious family emergency, military obligations, religious holidays). They should be communicated to the instructor before the missed class day when possible. UF rules require attendance during the first two

course sessions. Missing more than **three scheduled** sessions will result in a failure. Regardless of attendance, students are responsible for all material presented in class and meeting the scheduled due dates for class assignments.

Grading scale:

Letter Grade	Grade Points	Grade Percentage
A	4.0	95-100
A-	3.67	90-94
B+	3.33	87-89
B	3.0	83-86
B-	2.67	80-82
C+	2.33	77-79
C	2.0	73-76
C-	1.67	70-72
D+	1.33	67-69
D	1.0	63-66
D-	.67	60-62
E	0	< 59

For more details on letter grades and related University of Florida policies, please see the Grades and Grading Policies at <http://gradcatalog.ufl.edu/content.php?catoid=6&navoid=1219#grades>.

Make-up policy: Students are allowed to make up work only due to illness or other unanticipated circumstances. In the event of such an emergency, documentation will be required in conformance with University policy. Work missed for any other reason will earn a grade of zero.

UF POLICIES:

University policy on accommodation students with disabilities: Students requesting accommodation for disabilities must first register with the Dean of Students Office (<http://www.dso.ufl.edu/drc/>). The Dean of Students Office will provide documentation to the student who must then provide this documentation to the instructor when requesting accommodation. You must submit this documentation prior to submitting assignments or taking the quizzes or exams. Accommodations are not retroactive, therefore, students should contact the office as soon as possible in the term for which they are seeking accommodations.

University policy on academic misconduct: Academic honesty and integrity are fundamental values of the University community. Students should ensure they understand the UF Student Honor Code at <http://www.dso.ufl.edu/students.php>. You are expected and required to comply with the University's academic honesty policy (University of Florida Rules 6C1-4.017 Student Affairs: Academic Honesty Guidelines, available at <http://regulations.ufl.edu/chapter4/4017.pdf>). Cheating, plagiarism, and other forms of academic dishonesty will not be tolerated. Note that misrepresentation of the truth for academic gain (e.g., misrepresenting your personal circumstances to get special consideration) constitutes cheating under the University of Florida Academic Honesty Guidelines.

Netiquette – communication courtesy: All members of the class are expected to follow rules of common courtesy in all email messages, threaded discussions, and chats. The first instance of clearly rude and/or inappropriate behavior will result in a warning. The second instance will result in a deduction of five

percentage points from your overall grade. The third instance will result in a drop of a letter grade (A to B, A- to B-, and so on).

GETTING HELP:

For issues with technical difficulties for E-learning in Sakai, please contact the UF Help Desk at:

- learning-support@ufl.edu
- (352) 392-HELP - select option 2
- <https://lss.at.ufl.edu/help.shtml>

COURSE SCHEDULE (TENTATIVE):

The course schedule is subject to change based on our progress in each topic and students' backgrounds.

	Topic	Date	Notes
1	Introduction and course overview Introduction to clinical data science; definition, history, context, examples, technology landscape	8/27	
2	Essentials of Python programing: python grammars, variables, basic types, control structures, handle files. Hand-on sessions.	9/3	Bring your laptop
3	Database: database systems, common database models, relational database, entity-relationship model, E/R diagram, common database systems	9/10	
4	<i>EHR data</i> : structured/unstructured data, electronic health records systems, common data models, EHR data networks. Introduction to structured query language (SQL)	9/17	Bring your laptop
5	SQL practice + Guest lecture about statistical analysis	9/24	Guest lecture (Dr. Guo)
6	Machine learning part I	10/1	
7	Essentials of Python programing: do calculations, perform statistic summarizations, machine learning. Hand-on sessions.	10/8	Bring your laptop
8	Machine learning II + Guest lecture	10/15	Guest lecture (Dr. Xu)
9	Essentials of Python programing: do calculations, perform statistic summarizations, machine learning. Hand-on sessions. Introduction to course project	10/22	Bring your laptop
10	Midterm	10/29	
11	Natural language processing for medical text Midterm review Project QA	11/5	Guest lecture (Dr. Wu)
12	No Class - Attending AMIA	11/12	
13	Machine learning III	11/19	
14	Project QA	11/26	
15	Project QA	12/3	
16	Final Project Submission	12/10	