The Ohio State University College of Medicine, Department of Biomedical Informatics
BMI 5750 – Methods in Biomedical Informatics and Data Science
3 credit hours – Summer 2023

Course Directors: Amy Hite, PhD
Instructor E-mail Address: amy.hite@osumc.edu

Instructor's Office Hours: Wednesday 2-3pm
Class Time: Tuesday/Thursday, 1:30-3:05pm
Class Location: Cunz Hall room 140
Note: This course is hybrid with 25-74% in person and 25-74% synchronous zoom meetings. Of course, the total will be summing to 100%

Zoom link for classes and Office hours:
https://osu.zoom.us/j/99760643196?pwd=ckN3MEhKbm5Udm1DMXRCbnl3cGJ5Zz09
Meeting ID: 997 6064 3196
Password: 672525

Course description: This twelve-week course educates trainees in practical biomedical informatics, study design, statistical analysis, and computational techniques related to biomedical research. The course will provide applied primers covering foundational biomedical informatics and quantitative science methods employed in the design, conduct, and analysis of basic science, clinical, and translational research programs. This survey course is intended to enable individuals to critically select such methods and evaluate their results as part of both the design of new projects as well as the review of results available in the public domain (e.g., literature, public data sets, etc.). Core concepts to be reviewed during this course include:

1) Basic computational skills (R programming)
2) Data integration (data transformation / merging / manipulation, metadata integration)
3) Basic probability (conditional probability, probability distributions, sampling distributions)
4) Study design principles (population and sample selection, study design principles)
5) Exploratory analysis of data (graphical displays of data, data summarization)
6) Statistical analysis of data (estimation, confidence intervals, hypothesis testing, regression, two-group tests, analysis of variance (ANOVA), survival analysis)
7) In silico hypothesis generation (data mining, text mining, and visualization)
8) Introduction to data and methods in bioinformatics (clustering, classification, RNA-seq)

Class Format: Lectures and assignments.

Course Objectives: This course is designed for trainees in the biomedical sciences who have not received prior training in biomedical informatics, computational, or statistical theories/methods. The objectives of this course are to provide trainees with a solid understanding of core biomedical and quantitative science principles used in the design and execution of basic, clinical, and translational research programs. By the end of the class, trainees should be able to:

- Utilize R for common computational tasks and data manipulation
- Understand the basics of study design and how to appropriately obtain data for a research study

- Understand the basics of probability distributions and the population / sample paradigm
- Use methods to summarize and visualize data obtained from samples
- Implement common data analysis methods used in biomedical research
- Become familiar with common high-throughput data seen in bioinformatics (e.g., microarray and RNA-seq) and methods used to analyze this data
- Apply data mining for in silico hypothesis generation

Texts/Readings:
The primary material for the course will be the course lectures and associated articles / online resources. Additional readings, as assigned by the week's instructor(s), may be found on the Carmen site.

Grading:
During the course of the semester, students will be graded on:

Graded Item	**Contribution to Overall Grade**
Class Participation (each unexcused absence will result in a reduction of 2 points from the total possible 10% associated with class participation)	10%
Homework Assignments (3 total assignments, one per module, details will be posted to Carmen)	60%
Course Project (details will be posted to Carmen)	30%

Letter grades will be assigned using the following criteria (using a 100-point scale, based upon the preceding weighted axes): 100-93 = A; 92.9-90 = A-; 89.9-88 = B+; 87.9-84 = B, 83.9-80 = B-; 79-78= C+ 77-72 = C; 71-70 = C-; 69-60 = D; < 60 = Fail.

Assignments, Homework: Discussion write-ups and/or lab assignments will be an integral part of each educational module. The homework materials assigned each week are due for grading a week later (e.g. homework assignments distributed on Tuesday are due the following Tuesday) unless otherwise specified in class. Write-ups should incorporate theories and principles learned in the readings, lecture materials, and addressed in the in-class discussions. Any alterations to assignments or delivery due dates will be discussed in class and conveyed by email to the class.

Assignments, Final Project: The course will culminate with a final project. Students will work in teams of 3-4 members each (students wishing to do solo project should get approval from the course director). Each team is required to present a 20 minute oral presentation summarizing the findings and research. **Please note:** Final project presentations will take place during the last few weeks of the course. As soon as I know how many groups we will have, I will schedule the presentations in to our calendar. The final project is a group assignment with equal contributions from each student and we expect each student to present part of the material in person.

Course Materials: Carmen will be the primary delivery method for lecture notes, additional reading assignments, and guidelines for the final project. Readings will be posted two weeks in advance of classroom participation.

Class Policies: What you take away from this course will be a direct function of the effort you put forth inside and outside of class. While voluntary contribution is preferred, you can expect that you may be called upon at any time. After each class, the instructor will take notes on students' contributions to the class session. **If you do not attend class, it is impossible to receive credit for class participation.**

Effective class contribution entails providing good answers to questions. Effective comments add to our understanding of the underlying conceptual material, challenge, and clarify the ideas expressed by others, integrate material from past class work or other courses, and shows evidence of analysis rather than mere opinion or "gut feeling". Effective responses demonstrate that you have thought deeply about the material and can develop creative and innovative insights through this analytic effort.

Office of Disability Services

Any student who feels s/he may need an accommodation based on the impact of a disability should contact me privately to discuss your specific needs. Please contact the Office for Disability Services at 614-292-3307 in 150 Pomerene Hall to coordinate reasonable accommodations for students with documented disabilities.

Academic integrity: Academic integrity is essential to maintaining an environment that fosters excellence in teaching, research, and other educational and scholarly activities. Thus, The Ohio State University, the College of Public Health, and the Committee on Academic Misconduct (COAM) expect that all students have read and understood the University's *Code of Student Conduct* and the School's *Student Handbook*, and that all students will complete all academic and scholarly assignments with fairness and honesty. The *Code of Student Conduct* and other information on academic integrity and academic misconduct can be found at the COAM web pages (http://oaa.osu.edu/coam.html). Students must recognize that failure to follow the rules and guidelines established in the University's *Code of Student Conduct*, the *Student Handbook*, and in the syllabi for their courses may constitute "Academic Misconduct."

The Ohio State University's *Code of Student Conduct* (Section 3335-23-04) defines academic misconduct as: "Any activity that tends to compromise the academic integrity of the University, or subvert the educational process." Examples of academic misconduct include (but are not limited to) plagiarism, collusion (unauthorized collaboration), copying the work of another student, and possession of unauthorized materials during an examination. Please note that the use of material from the Internet without appropriate acknowledgement and complete citation is plagiarism just as it would be if the source were printed material. Further examples are found in the *Student Handbook*. Ignorance of the *Code of Student Conduct* and the *Student Handbook* is never considered an "excuse" for academic misconduct.

If I suspect a student of academic misconduct in a course, I am obligated by University Rules to report these suspicions to the University's Committee on Academic Misconduct. If COAM determines that the student has violated the University's *Code of Student Conduct* (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in the course and suspension or dismissal from the University. If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact me.

Course Outline

Weeks	Days	Topic(s)	Assignments
1	5/9-5/11	Course logistics & R install	
TOPIC 1: Introduction to R basics for biomedical research			
2-4	5/16-6/1	Data types, reading and writing data, data manipulation and merging, apply commands, writing R functions, basic R graphics, use of R Markdown to generate reports	Assignment #1 Due 6/2
TOPIC 2: Bioinformatics			
5-7	6/6-6/22	Gene expression (RNA-seq, arrays), differential expression, bioinformatics databases, multiple comparisons, clustering, classification, gene-set enrichment, basics of machine learning and data mining, The Cancer Genome Atlas (TCGA)	Assignment #2 Due 6/23.
TOPIC 3: Probability, statistics and data analysis			
8-10	6/27-7/13	Basic probability, probability distributions, sampling distributions, estimation, confidence intervals, hypothesis tests, ANOVA, linear models, survival analysis, logistic regression, overview of study designs (retrospective, prospective, clinical trials), reproducible research, power / sample size calculations	Assignment #3 Due 7/14
Final Project Presentations			
11-12	7/18-7/27	Final project presentations	