



## **MCB 447/547 – Big Data in Molecular Biology and Biomedicine Fall 2021: MWF 11-11:50, Harvill 332B**

### **Description of Course**

Recent technological advances enable the collection of massive biological data sets, both in the research lab and in the medical clinic. These Big Data offer opportunities for discovering new biology, but they also demand new analysis approaches. This course will introduce students with a strong molecular biology background to the use of Big Data statistics. Students will learn how to visualize complex data, identify biologically relevant clusters, model relationships between variables, and classify entities. They will apply these techniques to a diverse range of biological Big Data, including electronic medical records, gene expression measurements, and human population genetic sequences. Students will learn through homework, in-class exercises, and a substantive final project.

### **Course Prerequisites (for both 447 and 547)**

MCB 181R (Introductory Biology) or equivalent  
Math 263 (Introduction to Statistics and Biostatistics) or equivalent  
Math 119A, 122B, or Math 125 (Calculus) or equivalent

### **Course modality**

This class is scheduled to be taught in the in person modality.

### **Equipment and software requirement**

For this class you will need daily access to the following hardware: laptop; regular access to reliable internet signal; ability to download and run the following software: RStudio.

### **Instructor and Contact Information**

Prof. Ryan Gutenkunst, PhD

[rgutenk@arizona.edu](mailto:rgutenk@arizona.edu)

Office hour: Tuesdays at 1:00 pm, Life Sciences South 325

Prof. Megha Padi, PhD

[mpadi@arizona.edu](mailto:mpadi@arizona.edu)

Office hour: Mondays at 4:00 pm, <https://arizona.zoom.us/j/82747801479>, passcode: mcb447

The website for the course is on d2l, <https://d2l.arizona.edu/>.

### **Course Objectives and Expected Learning Outcomes**

In this course, students will:

- Learn data analysis techniques applicable to a wide range of biological Big Data, including visualization, clustering, regression, and classification
- Apply those techniques to a wide variety of biological Big Data in class and as homework
- Reproduce results from a scientific paper that applies Big Data techniques to molecular biology or biomedicine

After completing this class, students will be able to:

- Evaluate the suitability of various analysis approaches for different biological questions and data
- Use the programming language R to analyze biological data
- Visualize high-dimensional data and identify biologically-relevant clusters
- Model the relationship between a given outcome and multiple potential explanatory factors, including identifying the key factors
- Classify new data points using statistical models trained from existing data

Additionally, after completing this class graduate students will be able to:

- Evaluate the use of big data approaches in scientific literature relevant to their speciality

## Required Texts or Readings

*Machine Learning with R, the tidyverse, and mlr* by Hefin I. Rhys

Electronic access is available through the library, although we prefer a physical copy

## Assignments

Substantial class time will be spent on **in-class exercises**, often working in groups. If you are unable to finish the exercise in class, it will be due the following week. You are expected to ensure that all members of your group are following the work that is done. If you complete an assignment early, you are asked to join another group that has not finished and help out. Explaining your approach to others is an excellent way to learn material deeply.

**Preparatory homework** will be assigned often. These will be relatively straightforward applications of the approaches we will be exploring in class. The goal of these exercises is to give you experience working with the basic code implementing an approach, so that class time can be spent on more complex exercises requiring more interpretation.

**Summative homework** will be assigned several times throughout the semester. You are encouraged to work together and to discuss the homework assignments with your classmates. All submitted work, however, must be your own. Publications or online sources must be also cited.

**Extra credit** assignments will not be offered.

## Graduate Student Final Project

For graduate students, a final project for the course will center on applying big data approaches to real world biological or biomedical big data. The data may be either your own or from a published resource. In either case, you will analyze those data using techniques you have learned in the course. You will then give a **15-minute presentation** to your classmates explaining the scientific question, your approach, and your results. You will also turn in a **5-8 page paper** summarizing your scientific question, your approach, and your results.

To ensure you are using the skills you've developed in this course, your project must meet the following requirements:

- At least 50% of the analysis must be programmed by you in R or another equivalent language.
- If a web app is used for part of the analysis, there should be a short explanation of the algorithm underlying the web app in the presentation, and a longer mathematical description (with equations and possibly a derivation) provided in the paper. Limitations and assumptions underlying the web app should be clearly listed.
- All results in the paper/presentation should have a statistical evaluation (e.g. p-value) associated with them. If not, there should be a detailed justification for why it is not possible to evaluate the statistical significance, or suggestions for how it could be estimated in future work.

The final project presentations will be during the last week of the course, and the final paper will be due during finals week.

## Grading Scale and Policies

Final scores for the course will be calculated as below, and letter grades will be assigned based on the rubric below. If necessary, the instructors may lower the curve to shift the grade distribution upward.

Undergraduate students

Summative homework	50%
Preparative homework and In-class exercises	50%

Graduate students

Summative homework	35%
Preparative homework and In-class exercises	35%
Final project presentation	15%
Final project paper	15%

A	>90%
B	80-89%
C	70-79%
D	60-69%
E	<60%

**Late work policy:** Work less than one week late will have a 25% score reduction. Work more than one week late will have a 50% score reduction. These penalties can be waived with prior approval.

**Requests for incomplete (I) or withdrawal (W)** must be made in accordance with University policies, which are available at <http://catalog.arizona.edu/policy/grades-and-grading-system#incomplete> and <http://catalog.arizona.edu/policy/grades-and-grading-system#Withdrawal> respectively.

## Honors Credit

Undergraduate students wishing to contract this course for Honors Credit should email me to set up an appointment to discuss the terms of the contract. Information on Honors Contracts can be found at <https://www.honors.arizona.edu/honors-contracts>.

## Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities (e.g., texting, chatting, reading a newspaper, making phone calls, web surfing, etc.).

Students are asked to refrain from disruptive conversations with people sitting around them during lecture. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

## Absence and Class Participation Policy

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at: <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

The UA policy regarding absences for any sincerely held religious belief, observance or practice will be accommodated where reasonable, <http://policy.arizona.edu/human-resources/religious-accommodation-policy>.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: <https://deanofstudents.arizona.edu/absences>

Participating in the course and attending lectures and other course events are vital to the learning process. As such, attendance is required at all lectures and discussion section meetings. Absences may affect a student's final course grade. If you anticipate being absent, are unexpectedly absent, or are unable to participate in class online activities, please contact the instructor as soon as possible. To request a disability-related accommodation to this attendance policy, please contact the Disability Resource Center at (520) 621-3268 or [drc-info@email.arizona.edu](mailto:drc-info@email.arizona.edu). If you are experiencing unexpected barriers to your success in your courses, the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office is located in the Robert L. Nugent Building, room 100, or call 520-621-7057.

If you feel sick, or may have been in contact with someone who is infectious, stay home. Except for seeking medical care, avoid contact with others and do not travel.

Notify your instructor(s) if you will be missing a course meeting or an assignment deadline.

Non-attendance for any reason does **not** guarantee an automatic extension of due date or rescheduling of examinations/assessments.

Please communicate and coordinate any request directly with your instructor.

If you must miss the equivalent of more than one week of class, you should contact the Dean of Students Office [DOS-deanofstudents@email.arizona.edu](mailto:DOS-deanofstudents@email.arizona.edu) to share documentation about the challenges you are facing.

Voluntary, free, and convenient [COVID-19 testing](#) is available for students on Main Campus.

If you test positive for COVID-19 and you are participating in on-campus activities, you must report your results to Campus Health. To learn more about the process for reporting a positive test, visit the [Case Notification Protocol](#).

COVID-19 vaccine is available for all students at [Campus Health](#).

Visit the [UArizona COVID-19](#) page for regular updates.

## **Threatening Behavior Policy**

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See <http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>.

## **Academic advising**

If you have questions about your academic progress this semester, please reach out to your academic advisor (<https://advising.arizona.edu/advisors/major>). Contact the Advising Resource Center (<https://advising.arizona.edu/>) for all general advising questions and referral assistance. Call 520-626-8667 or email to [advising@arizona.edu](mailto:advising@arizona.edu)

## **Life challenges**

If you are experiencing unexpected barriers to your success in your courses, please note the Dean of Students Office is a central support resource for all students and may be helpful. The [Dean of Students Office](#) can be reached at (520) 621-2057 or [DOS-deanofstudents@email.arizona.edu](mailto:DOS-deanofstudents@email.arizona.edu).

## **Physical and mental-health challenges**

If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call (520) 621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.

## **Statement on compliance with COVID-19 mitigation guidelines**

As we enter the Fall semester, your and my health and safety remain the university's highest priority. To protect the health of everyone in this class, students are required to follow the university guidelines on COVID-19 mitigation. Please visit [www.covid19.arizona.edu](http://www.covid19.arizona.edu).

## **Threatening Behavior Policy**

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See <http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>.

## **Accessibility and Accommodations**

Accessibility and Accommodations: At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, <https://drc.arizona.edu>) to establish reasonable accommodations.

## **Code of Academic Integrity**

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: <http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity>.

The University Libraries have some excellent tips for avoiding plagiarism, available at <http://new.library.arizona.edu/research/citing/plagiarism>.

Selling class notes and/or other course materials to other students or to a third party for resale is not permitted without the instructor's express written consent. Violations to this and other course rules are subject to the Code of Academic Integrity and may result in course sanctions. Additionally, students

who use D2L or UA e-mail to sell or buy these copyrighted materials are subject to Code of Conduct Violations for misuse of student e-mail addresses. This conduct may also constitute copyright infringement.

## **UA Nondiscrimination and Anti-harassment Policy**

The University is committed to creating and maintaining an environment free of discrimination; see <http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>

Our classroom is a place where everyone is encouraged to express well-formed opinions and their reasons for those opinions. We also want to create a tolerant and open environment where such opinions can be expressed without resorting to bullying or discrimination of others.

## **Additional Resources for Students**

UA Academic policies and procedures are available at <http://catalog.arizona.edu/policies>

Student Assistance and Advocacy information is available at <http://deanofstudents.arizona.edu/student-assistance/students/student-assistance>

## **Confidentiality of Student Records**

<http://www.registrar.arizona.edu/personal-information/family-educational-rights-and-privacy-act-1974-ferpa?topic=ferpa>

## **Subject to Change Statement**

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructors.