BIO 331A Introduction to Bioinformatics   **SPRING 2021**

Instructor: Robert Harbert, Ph.D.
Contact Info: Email: rharbert@stonehill.edu
                   Office: SSC 204
                   Phone: 540-354-8104 (cell)
Office Hours: Tuesdays 8-9PM https://stonehill.zoom.us/j/93358217057
                   Or by appointment: sign-up here https://calendly.com/rharbert/scheduling

Communication: I want to make sure that you find it easy to reach out to me and that the course expectations are clear, organized, and easy to follow. The fastest way to reach me will be via text (using Slack) or via email (using your Stonehill address). You can, and should feel free to schedule an appointment with me any time using the Calendly scheduling app.

All the material related to this class will be posted in the eLearn course website. My hope is that you find it easy and straightforward to navigate. If you have any suggestions on how to make it better, please let me know.

Course Info: BIO331
        Course Website: https://bio331.devbioinformatics.org
        Course Slack: https://introbioinformatics.slack.com

Lecture Schedule: MW 4:00-5:15PM, SSC341 or
https://stonehill.zoom.us/j/97655904905?pwd=RVZpWWYybFBraWFoRCtna0Y5RHpIUT09

Lab: F 11:30-12:20, SSC341 or
https://stonehill.zoom.us/j/93358217057?pwd=bE5YbkN6RXFHNTNDK3M3TmllVEZ2dz09

Required Books:
"Bioinformatics Algorithms: An active learning approach" 3rd Edition, Compeau & Pevzner
ISBN: 9780990374633. https://www.bioinformaticsalgorithms.org/

(FREE ONLINE) "R For Data Science" Hadley Wickham and Garrett Grolemund. https://r4ds.had.co.nz/

BIO 331 Course Description:
This course introduces common concepts and tools in the field of Bioinformatics with a focus on developing a basic skill set for working with large biological data sets. The digital age has resulted in a period of rapid growth of data, and in biology this is revolutionizing how we look at the world.

Understanding how the field uses computational tools to manage and study these massive datasets is a crucial skill set for the modern Biology student. This course will cover the major sources of data in biology and an overview of the myriad of computational tools available.

Course Goals & Objectives:
After having completed Introduction to Bioinformatics you will be able to:
- Discuss Biology as a subdiscipline in "Data Science"

- Understand the major sources of 'big-data' in Biology and the scale and nature of the data being produced.
- Perform fundamental operations (data input/output, statistics, data visualization) in R.
- Use common bioinformatics tools through the Unix command line and coding in R
- Understand the importance of reproducibility and open access for data and computer code in bioinformatics.

BIO 331 Tentative Lecture Schedule

Date	Day	Lecture	Topic	Reading	Assignments DUE
1/24	M	1	Welcome		
1/27	W	2	Biology as a Data Science	Big Data: Astronomical or Genomical?	
1/29	F	L1			
2/1	M	3	Getting Started with R		Data Collection (Group)
2/3	W	4	More R	R for Data Sci	Big-Data in Biology (Group)
2/5	F	L2			
2/8	M	5	Data Visualization with R		
2/10	W	6	Intro to HPC		
2/12	F	L3			Data Visualization HW
2/15	M	7	Git and code management in Rstudio		
2/17	W	8	DNA Sequencing Technology		
2/19	F	L4			DNA Sequencing (Group)
2/22	M	9	Introduction to the Unix Shell		
2/24	W	10	Algorithms	Bias in Algorithms	
2/26	F	L5			Bash Script HW
3/1	M	11	Patterns in DNA sequence data – kmer counting		C&P Chapter 1 – Where in the Genome Does DNA Replication Begin?
3/3	W	12	Read mapping and Genome Variant Analysis		
3/5	F	L6			
3/8	M		**EXAM 1**		***After the exam this will be a "Slow Week". No other assignments will be due.

					Take a break while I catch up on grading!
3/10	W	*NO CLASS*	*Reading Day*		
3/12	F	L6			
3/15	M	13	Introduction to Genome Assembly		C&P Chapter 3 – How Do We Assemble Genomes?
3/17	W	14	Genome Assembly + Bash scripting		
3/19	F	L7	Assembly Lab – Long Read Genome Assembly		
3/22	M	15	Gene Prediction		
3/24	W	16	Open Science, bioRxiv, and course projects		
3/26	F	L8			
3/29	M	17	Sequence alignment and BLAST		C&P Chapter 5 – How Do We Compare Biological Sequences?
3/31	W	18	rBLAST	BLAST Paper (eLearn)	
4/2	F	*NO CLASS*			BLAST Homework
4/5	M	*NO CLASS*			
4/7	W	19	Metagenomics	Ashinnekoo, et al. 2015;	
4/9		L9			Metagenomics HW
4/12	M	20	Multiple Sequence Alignment		C&P Chapter 7 – Which Animal Gave Us SARS?
4/14	W	21	Phylogenetic Trees	Baum, et al. 2005;	
4/16	F	L10			
4/19	M		**EXAM 2**		
4/21	W	22	R for Data Science I – Organization and Communication		
4/23	F	L11			
4/26	M	23	R for Data Science II – Parallel Computing		
4/28	W	24	Project Peer Review		
4/30	F	L12			
5/3	M	25	R for Data Science III – More Data Vizualization		
5/5	W	26	Project Presentations		

TBA			**FINAL EXAM**	

Assignments:

Homework (40%): Will be assigned on a semi-weekly basis. These will typically involve extending and adapting the code and data analysis performed during class but using new datasets or additional functionality.

Discussions (20%): Each week there will be a discussion prompt related to the course material. Full credit requires active participation in the Slack discussion.

Exams (30%):

All exams are given as individual take-home exams with no-time-limit. Two mid-term and one final exam will be given in this class.

Mid-term Exams (2 @ 10% each)
Cumulative Final Exam (10%)

Term paper/project/presentation (10%): After the mid-term exam we will begin working on the term project for this course. Each student will identify an area of bioinformatics not covered in this course and do an exploratory project to investigate the kinds of analyses being done, the input data required, and the kinds of insights that can be derived from their chosen tool or method. Ideally, it will be possible to develop a brief teaching module where the chosen tool or method is demonstrated for the class. Unlike a typical term paper, this assignment will consist of a brief project report (1-2 pages single spaced) and a short (10 minute) presentation to be given in front of the class at the end of the term.

Late Assignments

Late work will be accepted with a penalty of 10% per day late. No assignment will be accepted more than 5 days after the assigned due-date (5 days late == 50% credit, 6 days late == 0%).

NOTE: If at any time during the semester you are having trouble keeping up with your work I am available to discuss alternate arrangements. I am happy to work with you on scheduling late assignments IF you communicate your needs.

Course Grade:

Your grade for BIO 331 will be calculated as follows:

Homework	40%
Discussion	20%
Exams	20%

Cumulative Final Exam	10%	
Term Project	10%	
Total	100%	

Grading Scale:

A	93.5 and above	A-	89.5 to 93.4		
B+	86.5 to 89.4	B	82.5 to 86.4	B-	79.5 to 82.4
C+	76.5 to 79.5	C	72.5 to 76.4	C-	69.5 to 72.4
D+	66.5 to 69.4	D	59.5 to 66.4	F	Below 59.5

Grade	Definition	Quality Points per Credit Hour
A	**Excellent**, work that is of the highest standard, showing distinction	4.00
A-		3.70
B+	**Good**, work that is of high quality	3.30
B		3.00
B-	**Satisfactory**, work that fulfills requirements in quality and quantity and meets acceptable standard for graduation	2.70
C+		2.30
C		2.00
C-	**Passing**, work that falls below graduation standard, yet is deserving of credit.	1.70
D+		1.30
D		1.00
F	**Failure**, work undeserving of credit	0.00

All grades will be posted in the course sections of eLearn. It is your responsibility to check the grades I have posted to make sure they match the assignments I have returned to you.

Academic Expectations:
BIO 331 is worth 3 credits, this means that **students need to schedule 3 hours of lecture time and a minimum of 6 hours of work/study time for BIO 331 each week**. This expectation is in agreement with the guidelines established by the Federal Government when they defined a college course credit. The official definition is shown below and was taken from the New England Association of Schools & Colleges, Inc. Commission on Institutions of Higher Education – 5th Year Interim Report Manual.

The Federal Government definition of **one course credit** is:
  "… an amount of work represented in intended learning outcomes and verified evidence of student achievement that is an institutional established equivalence that reasonably approximates not less than-

(1) One hour of classroom or direct faculty instruction and a minimum of two hours of class student work each week for approximately fifteen weeks for one semester or trimester hour of credit, or ten to twelve weeks for one quarter hour of credit, or the equivalent amount of work over a different amount of time.

<div align="center">OR</div>

(2) At least an equivalent amount of work as required in paragraph (1) of this definition for other academic activities as established by the institution including laboratory work, internships, practica, studio work, and other academic work leading to the award of credit hours."

Attendance, Punctuality and Courtesy:

Students are expected to attend all lecture and laboratory sessions and are responsible for all information covered during this time. Lecture material will be available online, but students will be responsible for obtaining class notes from a friend. Late arrivals can be disruptive to the class and are a waste of your tuition dollars.

Membership in this academic community places an obligation on all members to preserve an atmosphere that supports the freedom to teach and learn. Part of that obligation implies the responsibility to maintain a positive learning environment in which the behavior of any individual does not disrupt the class. It is the responsibility of the faculty to determine, maintain and enforce the standards of behavior acceptable to preserving an atmosphere appropriate for teaching and learning. Students will be warned and/or removed from class if their behavior is evaluated by the faculty member as disruptive according to policies outlined in the Student Handbook. Disruptive behavior may include, but is not limited to persistent speaking out without being called upon, refusing to be seated, or leaving the class early without instructor permission. Disagreement with the instructor or others is NOT, in itself, considered disruptive and the code of conduct will not be used to discourage classroom dissent.

Our classroom extends to our online discussions and interactions. For our online discussions and interactions please familiarize yourself with the Core Rules of Netiquette: http://www.albion.com/netiquette/corerules.html

Academic Honesty:
As a member of my class and the Stonehill community, it is expected that you adhere to the College's Academic Honor Code and Academic Integrity Policy. The College's official Academic policy can be found in the Hill Book. It states that "academic dishonesty includes but is not limited to giving or receiving, or attempting to give or receive, unauthorized assistance or information in an assignment or examination."

Any written work must be your own work (in this course "written work" INCLUDES computer code). Therefore, ideas and concepts should be in your own words and should not have similar sentence structure or wording to published work or another student's work. If you use a source to write a paper or lab report, you must cite this source. Simply altering a sentence from a source does not

make it your own work so be sure to cite. If you fail to cite the source, this is plagiarism. **Copying or sharing a computer file *at any stage of writing* will *absolutely not be tolerated*.** Such cases will be reported to the Dean of Academic Administration and failure of the entire course will ensue. Any electronic information or file that can be shared will be posted in the eLearn website by the lecture or lab instructor.

Violation of the academic policy can result, at my discretion, in either of the following penalties: (1) loss of credit in the exercise, or (2) failure in the course.

Inclusive Classroom Statement

Your experience and success in this class are important to me. All students in this class are treated equitably, have equal access to learning, and will feel valued and supported. In support of this, we will use various classroom strategies (teaching, assessment, and feedback) and include course materials that address the needs of students from a variety of backgrounds. Bioinformatics is an interdisciplinary subject that will require you to learn new analytical skills that will be challenging. Course material including code and background information will be available to you in-class during lectures, online with our course website, and through recordings of classroom demonstrations.

Commitment to an inclusive learning environment: Stonehill College embraces the diversity of students, faculty, and staff, honors the inherent dignity of each individual, and welcomes their unique cultural and religious experiences, beliefs, and perspectives. We all benefit from a diverse living and learning environment, and the sharing of differences in ideas, experiences, and beliefs help us shape our own perspectives. Course content and campus discussions will heighten your awareness to these differences.

There are many resources for anyone seeking support or with questions about diversity and inclusion at Stonehill. If you want more information on how to get connected to resources, the Office of Intercultural Affairs is a good first stop: Location: Duffy 149, Phone: 508-565-1409, Email: diversity@stonehill.edu.

If you are a witness to or experience an act of bias at Stonehill, you may submit a bias incident report online or on the Stonehill App. If you would like to learn more on bias incident prevention and response, or submit a report please visit: http://www.stonehill.edu/offices-services/intercultural-affairs/bias-response-protocol/

Resources for Academic Support

The Center for Writing and Academic Achievement (CWAA) provides academic support services in a welcoming, professional environment that emphasizes collaborative learning and peer tutoring, supplemented with professional-level support. The CWAA offers a variety of academic support services, including peer tutoring in writing, math, and foreign languages.

The CWAA is located in MacPháidín Library, Room 314. Drop-in hours are offered Sunday – Thursday. Students can visit the CWAA website to view schedules, make appointments, or request a tutor.

Cell Phone and Electronic Device Policy

It is understood that cell phone use is ubiquitous today. Minimal, non-disruptive cell phone use will be tolerated during lectures and practical activities. **Please keep cell phone use to a minimum, at no time should this detract from your class participation.**

Students with Disabilities

Stonehill College is committed to providing a welcoming, supportive and inclusive environment for students with disabilities. The Office of Accessibility Resources (OAR) provides a point of coordination, resources and support for students with disabilities and the campus community. If you anticipate or experience physical or academic barriers based on disability, please let me know so that we can discuss options. You are also welcome to contact OAR to begin this conversation or to establish reasonable accommodations for this or other courses. OAR is located within the Academic Services & Advising Suite in Duffy 104. For additional information please call (508) 565-1306 or email accessibility-resources@stonehill.edu.