Biology 364/664 Syllabus
Analysis and Visualization of Biological Data
Spring 2019
Tues Thurs 11:00am to 11:52am and Tues 1:00pm to 4:52pm

Prof. Ken Field

December 29, 2018

1 Contact Information

Prof. Ken Field

- 208 Biology Building
- kfield@bucknell.edu
- @ProfKenField on Twitter
- @KField-Bucknell on GitHub

2 Course Description

Introduction to the process of data exploration and visualization using state-of-the-art computational techniques. Using "big data" from public archives or their own research projects, students will learn how to rigorously analyze and visualize complex biological datasets. Lab will include hands-on work with R and virtual reality. No programming experience required.

3 Course Objectives

1. Students will analyze, visualize, and intepret real-world datasets using reproducible data science methods and R, R markdown, and Git.
2. Students will learn to identify and avoid questionable research practices when designing experiments, analyzing data, and presenting results.
3. Working as a team, students will complete novel projects utilizing whole-transcriptome or whole-genome datasets.
4. Students will present their final projects using complex multi-dimensional data visualizations.

4 Grading

$$\begin{aligned} Eight\,Homeworks,\ worth\,12.5\,points\,each &= 100pts. \\ Four\,Data\,Projects,\ worth\,25\,points\,each &= 100pts. \\ Takehome\,Midterm\,Exam &= 100pts. \\ Final\,Project &= 200pts. \\ Takehome\,Final\,Exam &= 100pts. \\ \textbf{Total} &= \textbf{600pts}. \end{aligned}$$

The Data Projects and Final Project will be graded using labor-based contract grading as described by Asao Inoue. These assignments will utilize a goal-oriented grading system as you develop your skills as a data scientist. When you or your group completes all of the goals associated with each project, you will earn 21/25 or 170/200 pts (85%). Failure to do so will result in a lower grade.

To earn more than 85%, you must do the following:

- Propose and perform additional analysis and visualization comparisons (discuss available options with me).
- Routinely assist in the learning and proficiency of your peers.

5 Textbooks and Readings

Required Texts and their Abbreviations, which are used on the syllabus:

- A Primer in Biological Data Analysis and Visualization Using R, by Gregg Hartvigsen, ISBN: 9780231166997 (Hart)
- R Programming for Data Science, by Roger D. Peng (PRP)
- Exploratory Data Analysis, by Roger D. Peng (EDA)

  These are Leanpub books, available from LeanPub for a "pay what you want" price.
- R for Data Science (R4DS) is a book teaching R for Tidy Data Science, and is available as a bookdown book on the web R for Data Science, you can buy an ebook for $18 from Google Play Store page for this book.
- Open Intro Statistics version 3 (OIS) is an open source text book on Inferential Statistics, published under a Creative Commons license, for free distribution as a pdf. In addition a copy can be purchased from Amazon for $9.
- Additional reading assignments will be distributed via the course git repository in the Readings subdirectory.

6 BIOL 364 Syllabus: Weekly Topics

Week:Date	Topic	Reading	Due
w1:15Jan2019	Using R, Rstudio, Git	HART1,2,3, PRP4,5	Making Graphs in R
w2:22Jan2019	Simple Data Exploration	HART4 PRP6,7,8,9,10 EDA4,5,6,7	Data Exploration in R
w3:29Jan2019	Data Visualization	HART5,11 EDA15,16 OIS4,5,6	**R Tutorial for 205 Project**
w4:5Feb2019	Advanded Hypothesis Testing	HART6,7,8 OIS7,8	Experimental Design
w5:12Feb2019	Questionable Data Practices	Fraser.pdf Frostmeier.pdf	Multiple Testing
w6:19Feb2019	Questionable Data Practices	Frostmeier.pdf	**QRP Case Studies Project**
w7:26Feb2019	Transcriptomics		QC & Mapping Statistics
w8:5Mar2019	Transcriptomics		**Transcriptomics Pipeline Project**
12Mar2019	**SPRING BREAK**		
w9:19Mar2019	Interactive Data	EDA5 minitufte.pdf	Data Archives **TAKEHOME MIDTERM**
w10:26Mar2019	Final Project		**Data Visualization Project**
w11:2Apr2019	Final Project	EDA12,13,14	Data Wrangling
w12:9Apr2019	Final Project		Preregistration
w13:16Apr2019	Final Project		**Final Project Presentations**
w14:23Apr2019	Final Project		**Final Project Presentations**
TBA	**TAKEHOME FINAL**		

Table 1: BIOL356/664 Weekly Syllabus. Peng R Programming for Data Science (PRPx), Peng Exploratory Data Analysis (EDAx), Open Intro Statistics (OISx) and A Primer in Biological Data Analysis and Visualization Using R (HARTx) refers to chapters assigned as reading.

7 Policies

7.1 Attendance

Your attendance at all classes and lab is expected, but not a graded part of the course. If you will need to miss lab for any reason, contact Prof. Field before class to make arrangements.

7.2 Readings

Readings must be done BEFORE the class where they are assigned. For dates with multiple reading assignments, browse each to determine the sections that are most useful to you. There will be overlap between the various textbooks and you should choose the book that is best for you and your individual background.

7.3 Homework Assignments

Homeworks are due before noon on Friday on the week they are assigned. Homework assignments will be submitted on GitHub (more instructions to follow). A 25% deduction will be assessed for submissions not received by noon on Friday. Assignments will not be accepated after noon on Monday.

7.4 Collaboration and Citation

For all projects and homework assignments working together is acceptable and encouraged. It is not ethical to do someone else's work or to have someone do your work. You must cite **all** resources used to work on your homework and projects. Citations should be done at the end of the document. These references can be to books, Wikipedia and other web resources, and discussions with other students. Working together and discussion is not allowed on takehome exams.

7.5 Academic Integrity Policy

Read Academic Responsibility at Bucknell for policies regarding academic integrity. Any questions concerning academic responsibility or misconduct will be referred to the Board of Review for Academic Responsibility without hesitation. Always cite the source of any information from outside sources, including online sources and classmates. Assignments may be screened using software designed to detect plaigarism. Unless explicitly directed otherwise, all takehome exams are expected to represent individual, not collaborative, work.

7.6 Bucknell University Honor Code

1. I will not lie, cheat, or steal in my academic endeavors.
2. I will forthrightly oppose each and every instance of academic dishonesty.
3. I will let my conscience guide my decision to communicate directly with any person or persons I believe to have been dishonest in academic work.
4. I will let my conscience guide my decision on reporting breaches of academic integrity to the appropriate faculty or deans.

7.7 Accommodations

Any student who needs an accommodation based on the impact of a disability should contact Heather Fowler, Director of the Office of Accessibility Resources at hf007@bucknell.edu, 570-577-1188 or in room 212 Carnegie Building who will coordinate reasonable accommodations for students with documented disabilities. The college will make reasonable accommodations for persons with disabilities.

8 License