

San José State University
Department of Biological Sciences
Biology 255E
Advanced Topics: Ecology and Evolution
Fall 2017

Course and Contact Information

Instructor:	Dr. Luke Miller
Office Location:	DH 336
Telephone:	(408) 924-4880
Email:	luke.miller@sjsu.edu
Office Hours:	Wed & Thurs 10:30-11:30
Class Days/Time:	Wed 14:30-17:20
Classroom:	Duncan Hall 344
Prerequisites:	Laptop and WiFi access

Faculty Web Page and MYSJSU Messaging

Course materials such as syllabus, handouts, notes, assignment instructions, etc. can be found on [Canvas Learning Management System course login website](http://sjsu.instructure.com) at <http://sjsu.instructure.com>. You are responsible for regularly checking with the messaging system through [MySJSU](http://my.sjsu.edu) at <http://my.sjsu.edu> (or other communication system as indicated by the instructor) to learn of any updates.

Course Description

This course will cover a set of practical computing skills relevant to graduate studies in the biological sciences. Topics include: Obtaining and importing data from external sources, and formatting your own data. Text file handling, data munging + cleaning, correcting your colleagues' questionable organization methods. Regular expressions for handling textual data. Generating basic and advanced graphics suitable for posters and publication. Establishing a reproducible workflow and adhering to reproducible research philosophies. Version control and archiving. These skills are meant to be subfield-agnostic and should be useful in all facets of biological research.

Course Learning Outcomes (CLO)

After completing this course, students will:

- Understand how to import data from external sources, interpret data formats, and reformat as necessary for analysis
- Understand how to use programming methods to automate data ingestion and formatting steps, and to subset data programmatically.
- Generate exploratory graphical displays and publication-quality graphics from data.
- Be comfortable using reproducible research strategies to organize and archive their analysis workflow.
- Use version control and data backup strategies to avoid organizational tragedies and lost work.

Required Texts/Readings

Relevant papers and materials will be provided throughout the semester. Suggested (not required) readings from *Practical Computing for Biologists* (2011) S.H.D. Haddock & C.W. Dunn, Sinauer Associates Inc. ISBN 978-0-87893-391-4

Other technology requirements / equipment / material

- **Required:** A laptop running Windows, OSX, or Linux. Chromebooks will not be sufficient.
- **Required:** A copy of the **free** statistical software R 3.3 or newer, available from <http://cran.cnr.berkeley.edu/>. Choose the version (Mac, Linux, Windows) appropriate for your computer. We will run R through RStudio Desktop, a 2nd download. RStudio is available for **free** from <https://www.rstudio.com/>, for Mac, Linux, or Windows, but not Chromebooks, iOS or Android. **This software will be available on the lab computers**, but you may wish to install it on your personal computer as well.
- **Required:** A free text editor program. Example suggestions: for Windows, Notepad++, for OSX, TextWrangler, for Linux, gedit.

Course Requirements and Assignments

SJSU courses typically demand a minimum of forty-five hours spent for each unit of credit, about three hours per unit per week. This course is 3 units. This includes preparing for class by doing reading and assignments, coming to class and participating in class activities, and diligently searching for answers to the questions that plague you, using the web and texts. More details about student workload can be found in University Policy S12-3 at <http://www.sjsu.edu/senate/docs/S12-3.pdf>.

NOTE that University policy F69-24 at <http://www.sjsu.edu/senate/docs/F69-24.pdf> states that “Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading.”

Course requirements include completion of quizzes, homework, a final portfolio, and participation in in-class activities.

Final Portfolio guidelines

The final portfolio will be due during the final exam period.

The end product for the term will be a “portfolio” consisting of data and associated analysis code that constitute a reproducible unit that could be placed in an archive and used to recreate the analysis exactly in the future. A

final analysis document detailing the workflow and results will accompany the portfolio. The portfolio will need to consist of the following components:

- A data collection consisting of one or more files that will be analyzed
- A repository of code used to ingest and analyze the data collection
 - The code repository will be in a version control framework and will exhibit evidence of having been regularly updated with useful comments
 - Analysis code will be contained in a Markdown or Rnw document that contains both the code and written explanations of the workflow (i.e. materials and methods) and results of the analysis, all of which can be used to generate a final PDF or HTML document with embedded figures
 - Figures will be vector or raster format, with a design and size appropriate for a print publication or poster
- The combination of data and analysis code will be sufficient to run on the instructor's computer and exactly duplicate the final analysis document submitted with the portfolio.

Grading Information

Determination of Grades

Final letter grades are calculated based on the percentage of the total possible points earned, with the following scale:

A+: 98-100%	A: 93-97.9%	A-: 90-92.9%
B+: 88-89%	B: 83-87.9%	B-: 80-82.9%
C+: 78-79%	C: 73-77.9%	C-: 70-72.9%
D+: 68-69%	D: 63-67.9%	D-: 60-62.9%
F: Below 60%		

Percentages will be rounded to one decimal place. There will be no grading curve or extra credit. As a result, you are not competing against your classmates for a certain letter grade; you are only striving to earn the best grade you can. Start earning full credit on the easy lab exercises early, and keep with it. The points for this course are apportioned as follows:

Quizzes	20
Homework	100
Final Portfolio	80
Total	200

If you believe that there is an error in the grading of an assignment or exam, you must submit a written request explaining why you believe there was an error in the grading, citing evidence from lecture, lab handouts, or the text, within one week of receiving the graded assignment or exam. All error correction requests will give the instructor the opportunity to ensure that no other possible errors or ambiguous answers were missed during the grading process, to the possible detriment of the final score.

Legalese: "All students have the right, within a reasonable time, to know their academic scores, to review their grade-dependent work, and to be provided with explanations for the determination of their course grades." See University Policy F13-1 at <http://www.sjsu.edu/senate/docs/F13-1.pdf> for more details.

Classroom Protocol

Lecture activities

Students are expected to attend lectures, to take notes, and participate as required.

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' [Syllabus Information web page](http://www.sjsu.edu/gup/syllabusinfo/) at <http://www.sjsu.edu/gup/syllabusinfo/>

BIOL 255E Advanced Topics

FALL 2017 Course Schedule

Subject to change with fair warning

Week	Date	Topics		Readings
1	Aug 23	Course Overview		R-intro
2	Aug 30	R basics Importing data		Reinhart-Rogoff https://goo.gl/qjiUFZ
3	Sep 6	Importing data II	Homework – importing data	
4	Sep 13	Reproducible research	Homework – importing data	Baggerly 2010 video Clark 2017
5	Sep 20	Markdown / LaTeX	Homework – importing data	
6	Sep 27	Version control	Homework – establish repo	Lowndes et al. 2017
7	Oct 4	Automating analyses		Wilson et al 2014
8	Oct 11	Regex, filtering, subsetting		Ziemann et al 2016
9	Oct 18	Regex, filtering, subsetting	Homework – filtering and subsetting	
10	Oct 25	Visualization I		Rougier et al. 2014
11	Nov 1	Visualization II	Homework – generating figures	Haddock & Dunn
12	Nov 8	Visualization III	Homework – generating figures	Weissgerber et al. 2015
13	Nov 15	Automating data collection		
14	Nov 22	Turkey Day (no class)		
15	Nov 29	Automating data collection		
16	Dec 6	Portfolio work		