Semester:	**Spring 2023**
Course Number:	**MAR 536**
Course Title:	**Biological Statistics II**
	Instructor:
	Gavin Fay, Associate Professor
	School for Marine Science & Technology
	SMAST East 228; (508) 910-6363; gfay@umassd.edu
	Lab teaching assistant:
	Lucy McGinnis (lmcginnis@umassd.edu)
Class Location:	SMAST East Rooms 102 & 103
	836 S Rodney French Blvd, New Bedford, MA 02744
	and synchronous via Zoom
Class Time:	Tuesday/Thursday 10:30-11:30
	Wednesday 10:30-12:00
Course website:	gavinfay.github.io/mar536-biolstats2
Student Hours:	By appointment (via Office365 Bookings)

COVID-19: We are in the middle of a global pandemic and things are not normal. I want you to succeed in this course and will do what I can to help support your being able to meet our learning objectives. Please communicate with me if there are additional things I can do to help support you. I will not ask you to disclose information related to your reasons for such requests outside of what is needed to provide possible solutions.

Code of Conduct: This class follows the Fay Lab Code of Conduct, which applies to all class spaces, including classrooms, Zoom calls, class GoogleDocs or other materials, and student hours. You can also read my lab's lab culture and philosophy, which underpins my approach to teaching and collaboration.

Course Description: Statistical analysis of ecological data. This course provides guided learning in advanced statistical analysis, as applied to ecological research and other fields of marine science. Advanced concepts in probability, statistical model building, hypothesis testing, inference, estimation and prediction, including analyses of example data sets and problems. Students will be required to learn and present topics to the class that are relevant to their graduate research and complete a class project. Course materials will be presented using the statistical programming language R; a computer lab focused on introductory and intermediate R programming will be held for the first half of the semester.

Course Objectives:

1. Self-learning in advanced quantitative concepts and methods
2. Familiarity with advanced statistical methods
3. Experience communicating advanced quantitative topics
4. Experience in statistical programming

Credits: 4

Prerequisites: Students should have taken an introductory graduate statistics course (e.g. MAR 535 Biological Statistics I or equivalent), or seek permission from the instructor.

Evaluation procedures:

1. An advanced statistical analysis of data relevant to the student's graduate research topic, including a detailed report and oral presentation (55% of course grade).
   a. 1-2 sentence **project topic idea due Jan 26** (1% of course grade).
   b. Brief (1-2 page) **description of the topic** including a statement of the problem, proposed statistical method, and supplemental reference(s) is **due Feb 09** (4% of course grade).
   c. **Data exploration report** containing data summaries and visualizations of the dataset to be used in the project, including observations of implications of data structure for the analyses to be conducted. **Due Feb 28**. (10% of course grade).
   d. **Student project presentation with chapter assignment or supplemental reading on the statistical method used in the project. A 10-minute lecture on the method, and demonstration of the method using the project dataset.** Presentations will be graded based on the instructors' evaluation of accuracy, clarity, and comprehensiveness. Feedback from other students will also be considered in evaluation of presentations (20% of course grade).
   e. Project report (**due Apr 25**). A written description of the methods and results of the project in the form of a draft scientific manuscript (Introduction, Methods, Results, Discussion). (20% of course grade).
2. **Mid-term examination on statistical concepts and literacy (20% of course grade). Take home exam. Exam distributed Feb 28. Due Mar 16.**
3. **Participation in class discussions** (5% of course grade). Attendance at all lectures and labs is the best way to understand topics and assignments, but is not required for evaluation.
4. **Computer lab exercises** (20% of course grade). Completion of short laboratory exercises using R. Lab exercises are due before lab session the following week. Individual weekly exercises will not be assigned numeric grades, assessment of the R lab component will take the form of i) completion of the exercises and ii) two self-reflection statements in the middle and at the end of the course asking students to assess their learning in R and ability to use these tools to address data-related and scientific questions. Additionally, the instructor and teaching assistant will provide general and specific feedback on R lab components and coding.
5. Extra credit: An additional 1% towards the total course grade can be obtained by students contributing to the #TidyTuesday, a public weekly data project in R from the R4DS online learning community (https://github.com/rfordatascience/tidytuesday). To obtain course credit, students should send Gavin Fay and Lucy McGinnis an email containing a visualization of analysis of one Tidy Tuesday data set, and a link to a repository containing code that reproduces the analysis and visualization (sharing of work with the #RStats community via Twitter or Mastodon is encouraged but not required for course credit).

Class Policies:

1. Grades will be reduced for all unexcused late submissions of assignments, with a 10% reduction for each day after the deadline. Requests for extended deadlines will be considered up to one day before the deadline. Unexcused assignments submitted later than three days after the deadline will not be graded.

2. Failure to complete any of the requirements for evaluation will result in a score of zero for missing components. At the student's request, and no more than 48 hours after the final exam or class, an incomplete grade may be given only in exceptional circumstances at the discretion of the instructor. The student must be passing the course at the time of the request or be sufficiently close to passing that the instructor believes that upon completion of the work, the student will pass the course. If the work is not completed within one year of recording the I, the grade will become an F(I).
3. This 4-credit class involves four hours per week of in-class work and an average of eight hours per week of out-of-class work is expected.
4. Assignment submissions should be made via a myCourses website. Students are responsible for all official correspondence sent to their umassd.edu e-mail address.
5. This course, including student participation, will be recorded on video by the instructor and will be available to students in the course for viewing remotely and after each session. As the Zoom recordings will contain student work, please do not share videos beyond the course attendees without the explicit permission of the instructor. Other course materials will be publicly available online.
6. If you have read this far, please use google search to find a picture or gif of an animal that best reflects your opinion of statistics, and send it to Gavin Fay attached to an email with the subject line "Here is a statistical opinion", worth an extra 5 points on the midterm.
7. University policy on academic dishonesty, including plagiarism, applies (see: https://www.umassd.edu/studentaffairs/studenthandbook/academic-regulations-and-procedures/).
8. Available academic support services are available at: www.umassd.edu/nfi/teaching-and-advising/course-syllabus/sample-disability-statement
9. SMAST Code of Conduct and Diversity statement: https://www.umassd.edu/media/umassdartmouth/smast/lab-pdf-files/SMAST-Code-of-Conduct-and-Diversity-Statement.pdf

A full description of Academic Policies associated with this and other UMass Dartmouth courses can be found at: https://www.umassd.edu/provost/resourcesforfaculty/syllabus-language/

Required Hardware: Class will take place in a SMAST-East classroom; individual laptop computers will be required to complete in-class and in-lab exercises and coursework. Pair programming approaches are encouraged during class sessions. Please contact the instructor if the computing requirement creates difficulty for your participation in the course; SMAST laptops may be able to be made available for in-class use. The university library also has laptops available for students to borrow.

Required Software: R (free download at http://r-project.org, students should also install Rstudio, an integrated development environment for R, free download at https://posit.co/download/rstudio-desktop/). I will walk through installation and use of this software during our first lab session. It is recommended to update your version of both R and RStudio for the course (I will be using R version 4.2.2). Web browser access to R sessions will also be provided through a cloud-based instance of RStudio via posit.cloud.

Course Materials: Materials will be distributed through a dedicated course website and RStudio Cloud project. All lectures and labs will be recorded using Zoom.

Remote participation: To facilitate off-campus attendance and engagement in the course, as well as ensuring course is available and accessible, class sessions will be broadcast via Zoom in addition to in-person meetings:

- https://umassd.zoom.us/j/94869202182
- Meeting ID: 948 6920 2182
- Passcode: 749571
- One tap mobile: +13052241968,,94869202182#,,,,*749571#

Principal Text:

Zuur, A.F., Ieno, E.N. and Smith, G.M. (2007). Analysing Ecological Data. Springer. 700p. Series: Statistics for Biology and Health. (available free ebook available through the UMass Amherst library system, or hard copy ~$100 online).
Support website for book (http://highstat.com/index.php/analysing-ecological-data)

The form for students to request access to UMass Amherst ebook collection is available at: https://www.umass.edu/it/accounts/courtesy-accounts-non-employees-non-students-nens

Supplementary Texts (others as needed):

Bolker, B.M. 2008. Ecological Models and Data in R. Princeton University Press. (http://ms.mcmaster.ca/~bolker/emdbook/index.html, http://emdbolker.wikidot.com/)

Ismay, C. and Kim, A.Y., 2019. Statistical Inference via Data Science: A ModernDive into R and the Tidyverse. CRC Press.

James, G., Wittem, D., Hastie, T., and Tibshirani, R. (2014). An Introduction to Statistical Learning With Applications in R. Springer. (ebook available online)
Support website for book (http://www-bcf.usc.edu/~gareth/ISL/)

McElreath, R., 2018. Statistical rethinking: A Bayesian course with examples in R and Stan. Chapman and Hall/CRC.

Wickham, H. and Grolemund, G., 2016. R for data science. O'Reilly. (http://r4ds.had.co.nz/)

Title IX statement: The purpose of a university is to disseminate information, as well as to explore a universe of ideas, to encourage diverse perspectives and robust expression, and to foster the development of critical and analytical thinking skills. In many classes, including this one, students and faculty examine and analyze challenging and controversial topics.
If a topic covered in this class triggers post-traumatic stress or other emotional distress, please discuss the matter with the professor or seek out confidential resources available from the Counseling Center, http://www.umassd.edu/counselling/, 508-999-8648 or -8650, or the Victim Advocate in the Center for Women, Gender and Sexuality, http://www.umassd.edu/sexualviolence/, 508-910-4584. In an emergency contact the Department of Public Safety at 508-999-9191 24 hrs./day.
UMass Dartmouth, following national guidance from the Office of Civil Rights, requires that faculty follow UMass Dartmouth policy as a "mandated reporter" of any disclosure of sexual harassment, abuse, and/or violence shared with the faculty member in person and/or via email. These disclosures include but are not limited to reports of sexual assault, relational abuse, relational/domestic violence, and stalking. While faculty are often able to help students locate

appropriate channels of assistance on campus, disclosure by the student to the faculty member requires that the faculty member inform the University's Title IX Coordinator in the Office of Diversity, Equity and Inclusion at 508-999-8008 to help ensure that the student's safety and welfare is being addressed, even if the student requests that the disclosure not be shared.
For confidential counseling support and assistance, please go to
http://www.umassd.edu/sexualviolence/

Course outline and schedule of lectures/labs: (subject to change)

Type	Day	Date	Reading	Topic	Instructor
Lecture	Tue.	Jan 17	Zuur et al. Chap. 1-4	Introduction, statistical rethinking	GF
Lab	Wed.	Jan 18		R Lab 1, Intro to R, working with data	GF
Lecture	Thu.	Jan 19	Bolker 2008 Chap. 4	Probability review	GF
Lecture	Tue.	Jan 24	Zuur et al. Chap. 1-4	data exploration, checking	GF
Lab	Wed.	Jan 25		R Lab 2, Visualizing data	GF
Lecture	Thu.	Jan 26	Zuur et al. Chapter 5	Linear regression review	GF
Lecture	Tue.	Jan 31	Hilborn & Mangel Chap 7	Likelihood	GF
Lab	Wed.	Feb 01		R Lab 3, probability & linear modeling	GF
Lecture	Thu.	Feb 02	Zuur et al. Section 6.1	Extending the linear model (GLM) (Poisson)	GF
Lecture	Tue.	Feb 07	Zuur et al. Section 6.2	GLM 2 (logistic regression)	GF
Lab	Wed.	Feb 08		R Lab 4, data wrangling & model summaries	GF
Lecture	Thu.	Feb 09	James et al. Chapter 5	Resampling methods, Cross-Validation	GF
Lecture	Tue.	Feb 14	James et al. Chapter 7	Nonlinear models, splines	GF
Lab	Wed.	Feb 15		R Lab 5, Iteration	GF
Lecture	Thu.	Feb 16	Zuur et al. Chapter 7	GAMs	GF
	Tue.	Feb 21		No class - Monday Schedule	
Lab	Wed.	Feb 22		R Lab 6, functions	GF
Lecture	Thu.	Feb 23	Zuur et al. Chapter 12	Matrix Algebra Review	GF
Lecture	Tue.	Feb 28	Zuur et al. Chapter 12	Principal Components Analysis **Midterm distributed**	GF
Lab	Wed.	Mar 01		R Lab 7, simulation	GF
Lecture	Thu.	Mar 02	Zuur et al. Chapter 14	Classification / Linear Discriminant Analysis	GF
				SPRING BREAK	
	Tue.	Mar 14	Zuur et al. Chapter 8	Linear mixed effects models	GF
	Wed.	Mar 15		R Lab 8, {brms}	GF
	Thu.	Mar 16		More mixed effects **Midterm exam due**	GF
Lecture	Tue.	Mar 21		Zero-inflated models	GF
	Wed.	Mar 22		UMass Intercampus Marine Science Symposium	
Lecture	Thu.	Mar 23		Spatial GLMMs	GF
Lecture	Tue.	Mar 28	Zuur et al. Chapter 9	Trees	GF
	Wed.	Mar 29		R Lab 9 (working withspatial data or tidymodeling)	GF
Lecture	Thu.	Mar 30		Penalized regression	GF
Lecture	Tue.	Apr 04	Zuur et al. Chap. 11-15	Ordination methods	GF
Lab	Wed.	Apr 05	James et al. Section 10.3	Cluster Analysis	GF
Lecture	Thu.	Apr 06		Presentations:	student
Lecture	Tue.	Apr 11		Presentations:	student
Lab	Wed.	Apr 12		Presentations:	student
Lecture	Thu.	Apr 13		Course Review	GF
	Tue.	Apr 18		No class – ECCWO5 meeting	
	Wed.	Apr 19		No class - Monday Schedule	
	Thu.	Apr 20		No class – ECCWO5 meeting	
	Tue.	Apr 25		**Project report due**	