BIO 6490 – Applied Biological Data Analysis
Course Syllabus and Schedule
Kennesaw State University – Fall 2023
Instructor: Nicholas S. Green, Ph.D.

1. Basic information

Section	Location	Day(s)	Time (US Eastern)	Modality
BIOL 6490	Burruss 468	MWF	1:25 – 2:15 PM	Face to face

2. Communication

2.1. Contact info

- **Email:** ngreen62@kennesaw.edu (best contact method). Please allow **1 business day** for a response.
- **Phone:** (470) 578-6546 (not a good contact method)
- **Messaging:** Microsoft Teams (KSU)
- **Office:** Science Building (SB) 331. We can also meet virtually using Microsoft Teams.
- **Office hours:** M 10:00 AM – 12:00 PM; T 9:00 – 11:00 AM; other times by appointment

2.2. Course communication

- **D2L** is used extensively for content delivery (e.g., lecture slides, datasets, etc.).
- Grades or other information covered by the Family Educational Rights and Privacy Act (FERPA) can only be communicated using official KSU channels (email or Teams associated with your KSU account) or face-to-face.

3. Course description

Prerequisites: *Statistical Methods I (STAT 3120) (3 hours) OR Biostatistics (STAT 3125) (3 hours) OR equivalent OR admission to the Masters of Science in Integrative Biology (MSIB) program.*
This course is a survey of data analysis skills and statistical methods that are essential for modern biology. The course takes a holistic approach to the data analysis workflow in biology using the open-source environment R, including data management, exploratory data analysis, data modeling, and reproducible science practices. Statistical topics covered include generalized linear models, mixed effects models, non-linear models, and ordination. Students are required to apply techniques learned in class to real or simulated biological datasets as a course project.

4. Course objectives

1. Explain the role of statistics in the biological sciences and the ways in which analytical results are communicated.
2. Manipulate, summarize, display, and analyze data using the open-source environment and language R.
3. Use probability distributions to model and think about biological phenomena. Students should come away from the course able to translate biological hypotheses and ideas into statistical statements and vice versa.
4. Conduct exploratory data analyses in support of scientific investigations, particularly to detect and diagnose common problems with biological datasets.
5. Employ modern statistical methods to answer biological questions.
6. Communicate data and analytical results to audiences who may or may not have statistical backgrounds.

5. Textbooks and other materials

5.1. Recommended texts

There are no required texts for this course. The following books are provided as recommendations for references on R or statistics. Much of the course is built around Bolker (2008).

Bolker BM. 2008. Ecological models and data in R. Princeton University Press, Princeton, New Jersey, USA. *General statistics textbook and introduction to R. Applications are focused on ecology, but the statistical exposition is relevant to any area of biology.*

Dalgaard P. 2008. Introductory statistics with R. Springer Science+Business Media, New York. *General purpose introduction to statistics and to R. Suitable for an introductory stats course, but also useful as a reference for R users of many levels.*

Illowsky B, Dean S. 2018. Introductory statistics. Rice University, Houston, Texas, USA. (Available free online). *Textbook designed for a college-level introductory statistics course.*

James G, Witten D, Hastie T, Tibshirani R. 2013. An introduction to statistical learning with applications in R. Springer Science+Business Media, New York. (7th edition available free online). A*dvanced general statistics textbook that includes both classical methods and more modern techniques such as machine learning.*

Zuur AF, Ieno EN, Smith GM. 2007. Analysing ecological data. Springer Science+Business Media, New York. *An advanced statistics textbook that, while focused on ecological applications, is suitable for any biologist.*

5.2. Technology requirements

Access to a computer or laptop capable of running R version ≥4.0.1. A 64-bit system is highly recommended. The classroom for this course, Burruss 468, is equipped with desktop computers for student use.

6. Evaluation and grading

6.1. General information

- Final letter grades are based on performance on exams, homework assignments, and the course project.
- Part of each exam will include a rubric so you can assess your own achievement of the course learning objectives. I will take your assessment into consideration when calculating your grade.
- Final course grades will be based on the traditional 10% grading scheme (≥90% = A, 80 – 89% = B, etc.), with percentages rounded to the nearest integer. Per KSU policy there are no +/- grades.

6.2. Grade calculation

Item	% of grade	Description
Homework	20	Periodic homework assignments (approximately biweekly) are used to assess student comprehension and to provide opportunity for practicing key concepts and procedures. Homework is graded on a completion basis. Collaboration between students is ***allowed and encouraged***.
Midterm exam	15	These exams are used to evaluate student understanding and progress at the middle and end of the semester. Exams are take-home, open book, and open note. Collaboration between students is ***not permitted***
Final exam	15	
Course project	50	The most important assignment of the course is a project where students must apply one or more of the statistical methods learned in class to a research question and dataset of their own choosing. A detailed assignment guide with project requirements and grading rubric will be provided early in the semester.

6.3. Attendance policies

6.3.1. Attendance

- Attendance is not required but is highly encouraged. This means that "excused absences" apply to assignment deadlines rather than routine class meetings.
- If an illness or other issue will prevent you from meeting an assignment or exam deadline, please contact me as soon as possible ahead of time so we can make arrangements.

6.3.2. Missed exams and assignments

- Assignments must be turned in on time to be eligible for full credit. Late assignments will be accepted but with a score penalty: 20% for up to 24 hours late; 40% for up to 48 hours late, and so on. The late penalty is waived if you miss the deadline for an excused reason (e.g., death in the family, serious illness of self or a loved one, jury duty, or other situations at the instructor's discretion).

7. Other course policies (instructor-specific)

- **Collegial and respectful behavior towards all people is expected.** This does not mean you cannot express your opinions; it means you must be respectful of other people.
- Children may accompany parents into the classroom, subject to the same behavioral expectations placed on students.

8. Academic honesty statement

Every KSU student is responsible for upholding the provisions of the Student Code of Conduct, as published in the Undergraduate and Graduate Catalogs. Section 5c of the Student Code of Conduct addresses the university's policy on academic honesty, including provisions regarding plagiarism and cheating, unauthorized access to university materials, misrepresentation/falsification of university records or academic work, malicious removal, retention, or destruction of library materials, malicious/intentional misuse of computer facilities and/or services, and misuse of student identification cards. Incidents of alleged academic misconduct will be handled through the established procedures of the Department of Student Conduct and Academic Integrity (SCAI), which includes either an "informal" resolution by a faculty member, resulting in a grade adjustment, or a formal hearing procedure, which may subject a student to the Code of Conduct's minimum one semester suspension requirement.

9. Institutional policies and resources

- KSU, BOR, and Federal policies applicable to this course are found at this link.
- Information on help and resources available to students can be found at this link.

10. Course schedule (tentative)

Week	Monday	Topic	Assignments and Exams
1	Aug 14	Syllabus; Intro to R	Pre-assessment due *Aug 20* @ 11:59 PM
2	Aug 21	Advanced R: Data manipulation and management	Project: Topic selection due Aug 27 @ 11:59 PM
3	Aug 28	Basic statistics review	Homework 1 due Sept 3 @ 11:59 PM
4	Sept 4	Exploratory data analysis 1: Data description and summarization	
5	Sept 11	Exploratory data analysis 2: Distributions and transformations	Project: Study plan due Sept 17 @ 11:59 PM
6	Sept 18	Exploratory data analysis 3: Exploring multiple variables	Homework 2 due *TBD* @ 11:59 PM
7	Sept 25	GLM 1: Log-linear and count models	
8	Oct 2	GLM 2: Proportional and binary models	Midterm due Oct 8 @ 11:59 PM
9	Oct 9	Nonlinear models	Project: Progress check 1 due Oct 13 @ 11:59 PM Homework 3 due Oct 15@ 11:59 PM
10	Oct 16	Mixed effects models	
11	Oct 23	Modeling review	Homework 4 due Oct 29 @ 11:59 PM
12	Oct 30	Multivariate 1: Dissimiliarity and ordination	
13	Nov 6	Multivariate 2: Multivariate analyses	Project: Progress check 2 due Nov 10 @ 11:59 PM
14	Nov 13	Wrap-up: planning and presenting your analysis	Homework **3** due Nov 19 @ 11:59 PM
15	Nov 20	*No class – Thanksgiving Break*	
16	Nov 27	Course project presentations	
17	Dec 4	Course project final deliverable due: Monday December 4, 11:59 PM Final exam due: Tuesday December 5, 11:59 PM EST	