Quantitative Biology - BIO 5324 Spring 2026 Syllabus

Location: BSB C230
Time: Tuesdays/Thursdays 9:30 - 10:45 AM

Instructor: Caleb J. Robbins, Ph.D.
Email: Caleb_Robbins@baylor.edu
Office: BSB C239.1
Office Hours: By Appointment.

Course Description: Quantitative approaches to biological datasets, introductory programming, management of data and workflows, visualization, exploratory analysis, communication of technical information, collaboration in teams, FAIR data principles, and data policies.

Course Prerequisites: Graduate student status or departmental and instructor approval.

Supplies: A working laptop will be necessary to follow lectures, do homework assignments, and take exams. Anything built within the last ten years should work (i.e., 4+ GB RAM, modern CPU) as there is no expectation for truly big or computationally intensive data analysis.

Learning Objectives
- Write code to wrangle, tidy, and visualize data.
- Collaborate using version control software, git, and Github.
- Understand the core concepts and relationships among study design, simulation, statistical modeling, and hypothesis testing.
- Understand and avoid common Questionable Research Practices common in data analysis in the biological sciences.
- Plan research projects centered around open science practices.

Course Texts

(CRH) Çetinkaya-Rundel, M and J. Hardin. 2025. Introduction to Modern Statistics. 2nd Ed. OpenIntro Statistics. Freely available: https://openintro-ims.netlify.app/

(IKV) Ismay, C., A.Y. Kim, A. Valdivia. 2025. Statistical Inference via Data Science: A ModernDive into R and the Tidyverse. 2nd Ed. CRC Press. Freely available: https://moderndive.com/v2/

Class Schedule

Date	Unit	Lecture	Reading	Notes
Jan 20	1: Learning the Lingo	Software Setup, and Why are we here?		
Jan 22	1: Learning the Lingo	The tools: Introduction to R and Github	IKV Ch. 1	PS 1 Due
Jan 27	1: Learning the Lingo	Wrangling your data (Yeehaw)	IKV Ch. 3 & 4	PS 2 Due
Jan 29	1: Learning the Lingo	A grammar of graphics	IKV Ch. 2, CRH Ch. 4 & 5	
Feb 03	1: Learning the Lingo	Dr. Chan-Park on Data Practices		PS 3 Due
Feb 05	1: Learning the Lingo	Data visualization for telling stories	CRH CH. 6	
Feb 10		Literature Presentations 1		
Feb 12		**Exam 1: Unit 1 (75 pts)**		PS 4 Due
Feb 17	2: Estimating quantitative relationships	What generates your data? Causes and models	CRH Ch. 2	
Feb 19	2: Estimating quantitative relationships	Quantifying relationships: Correlation and Linear models	CRH Ch. 7	
Feb 24	2: Estimating quantitative relationships	Quantifying (complicated) relationships: Multiple predictors	CRH Ch. 8	PS 5 Due
Feb 26	2: Estimating quantitative relationships	Quantifying relationships: Binary outcomes	CRH Ch. 9	
Mar 03		Literature Presentations 2		Recorded presentations; PS 6 Due
Mar 06	2: Estimating quantitative relationships	Quantifying (very complicated) relationships: A brief primer on GLMs, GAMs, and machine learning		Asynchronous lecture
Mar 17		**Exam 2: Unit 2 (50 pts)**		
Mar 19	3: Making decisions with samples	Foundations of inference: Empirical approach	CRH Ch. 11 & 12	Analysis Presentation Dataset Approval
Mar 24	3: Making decisions with samples	Foundations of inference: Theory-based approach	CRH Ch. 13	
Mar 31	3: Making decisions with samples	Hypothesis testing	CRH Ch. 14	

Apr 02	3: Making decisions with samples	Inference on single or two proportions	CRH Ch. 16 & 17	PS 7 Due
Apr 07	3: Making decisions with samples	Inference on two-way tables	CRH Ch. 18	
Apr 09	3: Making decisions with samples	Inference on single, two independent, or paired means	CRH Ch. 19, 20, & 21	PS 8 Due
Apr 14	3: Making decisions with samples	Inference for comparing many means (ANOVA)	CRH Ch. 22	
Apr 16	3: Making decisions with samples	Inference with Linear regression	CRH Ch. 24 & 25	PS 9 Due
Apr 21		Literature Presentations 3		Analysis Presentation Stats/Question Approval
Apr 23	4: Research workflows	Quantitatively-explicit study design		
Apr 28	4: Research workflows	Open to criticism: Data curation, reproducibility, and integrity		
Apr 30		Collaborate on Analysis Presentations		PS 10 Due
May 05		**Exam 3: Units 3&4 (75 pts)**		
May 09		**Final Exam: Analysis Presentations (100 pts)**		

Assessment

<u>Points Breakdown</u>
Class Participation: 40
Exams: 200
Problem Sets: 100
Literature Presentations: 60
Analysis Presentation: 100
Total: 500

<u>Grading Scale</u>
A: >90%
B: 80-89%
C: 70-79%
D: 60-69%
F: <60%

Class participation (40 pts)

Participation points are lost, not gained, for not actively participating in discussions, unexcused absences, or violating the Classroom Community Policy.

Problem Sets (100 pts)

Problem Sets (PS) will focus on practice skills and building intuition on concepts covered in lectures. The PS are given and completed in RMarkdown, allowing students to respond via both runnable code and written responses. Problem sets will be downloadable from the main Github repository https://github.com/robbinscalebj/QuantBio_BIO5324_Sp2026 and are due by class

time on the date listed. The timestamp of the git commit will be used for timing (students should make sure to Push these commits on time as well).

Completed PS will be turned in via the student's personal github repository for the class. Only the student and instructor have access to this repository. Problem set grades will not be contingent on whether the .Rmd renders, just that the code runs as necessary to answer the problems. The instructor will grade PS by cloning the student's repository, checking the code results and student responses in the .Rmd, and submitting an Issue to the student's repository with the grade for the PS. The student can discuss the grade through the discussion thread in the Issue.

Questions regarding problem sets (either noting errors or clarifications) or other course material can be collaborated on by the whole class by opening an Issue on the *main* Github Repository. Issues should clearly link to the Problem Set/question, if applicable, and be formatted as a reprex (Reproducible Example). A reprex should "just run" when others enter the code into R. For the purposes of this course, a reprex constitutes at least a chunk of code that loads the necessary libraries, reads or creates relevant data, and clearly shows the problem for which you are looking for help (e.g., spits an error clearly derived by running a specific piece of the code). Students who identify and/or solve errors in the problem sets will get bonus points on the problem set.

Problem Set	Topic	Due
PS 1	Software Setup	NA
PS 2	Introduction To R	NA
PS 3	Data Wrangling	NA
PS 4	Data Visualization	NA
PS 5	Causes, Models, And Simulation	NA
PS 6	Linear Modeling	NA
PS 7	Hypothesis Testing Foundations	NA
PS 8	Inference On Single Proportions And Means	NA
PS 9	Inference On Anova And Linear Regression	NA
PS 10	Design And Power Analysis	NA

Exams (200 pts)

Exams will cover material in readings, lectures, and assignments. They will be practical in nature, assessing coding and data analysis. Exams will also technically be cumulative, since skills learned early in the course will be foundational for skills later in the course.

Literature Presentations (60 pts)

Literature Presentations are designed to engage you with current problems in modern quantitative analysis. Pairs (and a group of 3 when the class has an odd number of participants) present to the class on an assigned paper. Material presented is assessable on exams and participation points may be deducted from audience members for inadequately engaging with presentations (e.g., demonstrably not paying attention).

Literature Presentation Rubric

	Full Credit (5 pts)	Half Credit (2.5 pts)	No Credit (0 pts)
Problem Explanation	The problem addressed by the paper is fully presented, including a brief explanation of the authors rationale and 'bigger picture', and linked to class material	The problem addressed by the paper is stated, but poorly contextualized to course material or the authors own rationale for writing the paper	Did not participate
Presentation of Results	Results from the paper are recapped (including showing a key graphic as necessary) to show how the authors answered the problem they were trying to address	Results are only briefly touched on	Did not participate
Reflection	Presenters share thoughts on the paper, which can include critical disagreement, related personal experiences that contextualize the importance (or not) of the results, views on utility of the results, etc.	Presenters share only light personal opinion but no critical thoughts	Did not participate
Discussion	Class is prompted with at least one thoughtful discussion question based on the paper	Discussion question is vague and uncritical, such as only asking for others vague personal opinion ("What do you think about that?")	Did not participate

Analysis Presentation (100 pts)

Each student will submit and present (5-7 min) a small data analysis as a final project, comprising a research workflow using the skills developed in the course. Data may come from any source, including student-generated from experiments, existing databases, or synthesized from the literature. However, the original data should be in a weakly or poorly tidied format (OK if tidied versions already exist), and data provenance will be clearly stated and traceable. The student will use the data to tell a short story (an analysis) linked to a scientific question/hypothesis (which does not need to be novel or interesting) within a broader scientific framework. The story should be told with both graphical techniques and a statistical analysis that derive from this question and the data. Presented discussion of the analysis will include the interpretation of

The dataset and statistical technique/question will be pre-approved by the instructor by the dates listed in the Syllabus Class Schedule Notes.

The Analysis Presentation does not need to be a rendered Analysis.Rmd, but material from the Analysis Presentation should be traceable to Analysis.Rmd and other files.

The entire analysis will be saved as a data package - really just a subdirectory in the student's Github repository. The folder structure should be as follows:

- Analysis Presentation/
  - Data/
    * *data.csv*
      · The original data file, basically unaltered. Any file format acceptable.
    * *data_description.csv*
      · This describes, with text, the variables in your data file. Columns are "Variable Name", "Description", "Units", "Type" (e.g., nominal, categorical), "Numeric Range" (when applicable), "Unique values" (when categorical)
    * *data_provenance.txt*
      · Describes who created the original dataset, when and where the data were collected (time and place of experiment or observations), and how the data were collected (methods).
    * *tidied-data.csv*
      · File format should be .csv
  - Analysis/
    * Figures/
      · Saved versions of central figures shown in the Analysis Presentation, at appropriate sizes in .png or .jpeg format.
    * *Download-and-Tidy.Rmd*
      · This Rmd shows your data provenance, taking raw data (data.csv) and converting it into transformed data (tidied-data.csv)
    * *Analysis.Rmd*
      · Records exploratory visualization, statistical analysis, interpretive text, and creates and saves key figures that are used in the Analysis Presentation.

Analysis Presentation Rubric

	10 pts	8 pts	4 pts	0 pts
Prep Work	Data set, analysis, and quantitative question chosen on time (see Syllabus). Data package is committed to github prior to presentation.	NA	NA	Late; Totally uncommitted data package by the time of presentations may result in 0 points across all Data Package subcriteria.
Data package: Metadata	Metadata files and original data all present, in the correct file format, and fully describe data provenenance (who, what, where, when collected the data; methods; URL source)	Metadata files and original data all present but some are missing information.	Metadata files and original data file only partially present, or files contain significant errors	Metadata files and original data are missing
Data package: Visualization	Saved figures are legible, with sensible choices for axis labels, font size, geom shapes/sizes/colors, etc., with conventional figure sizes. Captions are standalone descriptions of the figure, describing, e.g, any abbreviations, meaning of error bars, minor description of the data, etc.	Figures are minorly difficult to interpret graphically or caption only weakly describes the figure.	Figures do not coherently convey information about the data (e.g., unlabelled axes highly irregular tick spacing or labeling, aesthetic mappings not specified by either a legend or the caption).	Saved figures are missing
Data package: Tidy data	Tidied data used in the analysis are saved in a tidy and clean format (observations are rows, variable names and factor levels have consistent, meaningful naming convention; any times clearly have a time zone)	tidied_data.csv has minor non-compliances with tidy/clean formatting,	tidied_data.csv is "messy data" and only some portion of the data are clean	tidied_data.csv is a copy of the original data or missing

Data Package: Reproducibility	Both .Rmd files in Data package run on instructor's computer with no change to code, clearly creating a tidied data file and figures	.Rmd files require a minor alteration to run (e.g., load libraries, or absolute file paths used instead of relative file paths)	.Rmd files contain multiple errors requiring alteration for the instructor to run; or figures and tidy data are not created by the .Rmd files	.Rmd files are missing
Presentation: Introduction	Rationale provided for a precisely stated question/hypothesis and its link to the broader context of the scientific literature	Introductory material is lightly skimmed, but a question/hypothesis central to the analysis is presented	Introductory material is cursory and only tenuously connected to the question/hypothesis, or a question/hypothesis is not precisely stated	A question/hypothesis is not stated or is not contextualized to a broader scientific framework
Presentation: Data Description	Clear presentation of data collection/provenance and original study design (briefly highlighting metadata necessary to understand data and its relationship - even if poor - to the question/hypothesis)	Data are only briefly described, but the linkage to how they address the question/hypothesis and any study design is well described	Data are only briefly described and not well-linked to study design	Study design is not well described or Data are not described
Presentation: Statistical Analysis	Given the data and study design, clear rationale and results of statistical analysis answering the question/hypothesis, including identification of the null hypothesis and estimand.	Well-stated rationale, estimand, and null hypothesis, but statistical analysis results weakly presented	Statistical analysis weakly presented and not unclear rationale, estimand, or null hypothesis	Rationale, estimand, and null hypothesis are missing or no statistical results presented
Presentation: Visualization	Use of clear data visualization to tell story about an answer to the question/hypothesis, and to justify assumptions for statistical analysis	Data visualizations largely legible, but inappropriately sized or fuzzy	Data visualizations do not tell a clear story linked to the question/hypothesis, or not well described to the audience	Data visualizations very difficult to interpret or not connected to question/hypothesis or justifying analysis

Presentation: Discussion	Evidence from the results are used to justify inference about the question/hypothesis and linked back to the broader scientific framework; data/sample size and statistical assumptions considered in weighting evidence, particularly as presented graphically or with power analysis.	Question/hypothesis clearly answered by appealing to graph or results from statistical analysis, but considerations of data properties, study design/sample size, or statistical assumptions are not strongly considered (e.g., only rhetorically rather than quantitatively), or no linkage to broader scientific framework in Introduction	Robustness of analysis to answer the question/hypothesis is not considered	Question/hypothesis is not clearly answered.
Presentation: Length	<7 min	<8 min	<9 min	>10 min

*Note: The rubric has 110 points possible.

Policies

AI and LLM Policy

I will not use AI or LLMs to create or grade/evaluate any material for this class. This class will build intuition and skills that may facilitate the responsible use of LLMs in research later in your career, if you so choose. Dr. Andrew Heiss has some good thoughts and links to resources about AI in academics here. All assignments should be your original work and should not be produced in part or in total with the assistance of artificial intelligence (for example, ChatGPT, Grammarly, or some other resource). Students may not use note taking and content capturing tools in class without explicit permission. Use of artificial intelligence and other technologies without my explicit permission constitutes a violation of the Honor Code at Baylor University.

Attendance Policy

Attendance will be expected and counted in Class Participation points. Class participation indicates that the student is also engaged with the class activity, such as asking and answering questions or actively taking notes.

Absences

Absences for research activities (e.g., field work, time-sensitive lab work, conferences), religious observances, or personal and family emergencies should be discussed with the instructor as soon as possible prior to the absence but will typically be excusable and any in-class points can be made up out-of-class, for example, with a written report instead of the in-class group presentation.

Classroom Community

We learn best in welcoming environments and everyone deserves to be treated with respect and dignity in this classroom space: students, teaching assistants, instructors, and guests. Baylor's Code of Ethics (BU-PP 024) states, "Members of the Baylor University community are expected to act in a way that builds a distinctive sense of caring, kindness, mutual respect, collegiality, and fairness. The idea that the Baylor community is a family has always been vital in our historical traditions, and we here reaffirm our commitment to that idea. Out of respect for this community, we do not slander or defame." Incivility, unkindness, and aggression have no place in our classroom community. Such actions will result in consequences ranging from deduction of participation points to removal from the classroom and disciplinary referrals.

Academic Integrity

Academic dishonesty and any form of cheating involves a breach of student-teacher trust. Instances of plagiarism or any other act of academic dishonesty will be reported to Office of Academic Integrity, may be referred to the Honor Council for action, and may result in failure of the course. I expect you, as a Baylor student, to be intimately familiar with the Honor Code.

Academic and creative work submitted under your name is expected to be your own, neither composed by anyone else as a whole or in part, nor handed over to another party for complete or partial revision. Be sure to document all ideas that are not your own. Furthermore, in your completion of course work and assignments you must not use technologies which give you an unfair academic advantage for this course unless they have been explicitly permitted by me. In addition, providing course materials to other students, whether individually or generally (such as online) that would enable them to gain an unfair academic advantage is an act of academic dishonesty. If you have questions about what constitutes academic integrity in this course, please contact me.

Baylor Land Acknowledgement

We respectfully acknowledge that Baylor University in Waco and its original campus in Independence are on the land and territories originally occupied by Indigenous peoples including the Waco and Tawakoni of the Wichita and Affiliated Tribes, the Tonkawa, the Nʉmʉnʉʉ (Comanche), Karankawa, and Lipan Apache. These Indigenous peoples were dispossessed of and removed from their lands over centuries by European colonization and American expansionism. In recognition that these Native Nations are the original stewards of Baylor's campus locations, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.

Crisis and Emergency Numbers

Please make a note of the following numbers for crises or emergencies:

- Counseling Center Crisis Line: (254) 710-2467 (Business Hours/Non-Business Hours/Weekends)

- Baylor Police Department: (254) 710-2222

- MHMR Crisis Center: (254) 867-6550

- MHMR 24-Hour Emergency/Crisis Number: (254) 752-3451

- When home during academic breaks, when the counseling center is closed, please call your local resources and/or national hotlines:

    - National Hope Network Hotline: 1-800-SUICIDE (1-800-784-2433)

    - National Suicide Prevention Lifeline: 1-800-273-TALK (1-800-273-8255)

Accommodations

Any student who needs academic accommodations related to a documented disability should inform me immediately at the beginning of the semester. You are required to obtain appropriate documentation and information regarding your accommodations from the Office of Access and Learning Accommodation (OALA). OALA is located in the basement level of the east wing of the Sid Richardson building (reachable via elevator or stairs). They can also be reached by phone (254) 710-3605 or email OALA@baylor.edu.

Student Wellbeing

Students face many struggles. The Department of CARE Team Services is a group of dedicated and caring case managers who work with students facing mental health issues, financial struggles, and anything else affecting a student's ability to be successful in the classroom. You can find them on the second floor of the Student Life Center, suite 207. Contact Information: (254) 710-2100; CareTeam@baylor.edu.

Health Care
Baylor University is strongly committed to addressing the physical well-being and mental health of students by providing access to on-campus healthcare resources. Baylor Health Services includes Primary Care, Psychiatry, Physical Therapy and Pharmacy and is staffed with fully certified and licensed physicians and nurse practitioners, as well as nurses and administrative staff. Appointments may be made by calling their main number or by logging into the health portal located on their website. Contact Information: (254) 710-1010; Health_Services@baylor.edu.

Mental Health
You also have access to mental health services through The Counseling Center. Connect with them for an initial appointment, goal-directed individual therapy, group therapy, nutritional counseling, and case management services. You can schedule your 30-minute initial appointment online through the health portal or by calling (254) 710-2467. If you are experiencing suicidal thoughts, call or text #988 anytime day or night for support.

Telehealth
All Baylor students also have access to medical and counseling services via telehealth in addition to the on-campus services. Medical services include virtual urgent care, psychiatry, and nutrition counseling. Counseling services include 24/7/365 in-the-moment support and ongoing mental health counseling. Access to care is available in the evenings, during weekends, and when the University is closed. All services are free and unlimited to Baylor students. Please visit https://baylor.academiclivecare.com/ for more information.

Addiction
The Beauchamp Addiction Recovery Center (BARC) supports students in recovery from substance and behavioral addictions through an all-encompassing level of support approach that includes one-on-one mentorship, support groups, and social events open to all Baylor students. Located in the East Village Residential Community (bottom floor of Teal Residential College). Contact Information: (254) 710-7092; BARC@baylor.edu

Food Insecurity
At Baylor, we want all students to have access to food resources that will

support their holistic well-being and success. If you or someone you know experiences food insecurity at any time, you can find information on campus and community food resources by visiting The Store. You can also contact Store staff at (254) 710-4931. For additional basic needs assistance, please reach out to CASE or the Care Team.

Spirituality
The Office of Spiritual Life offers programs, persons, and resources to nurture theological depth, spiritual wholeness, and missional living. Located on the corner of 5th and Speight Street in the Bobo Spiritual Life Center. Contact Information: (254) 710-3517; Spiritual_Life@baylor.edu.

Sexual and Interpersonal Misconduct Policy and Civil Rights Policy

Baylor University does not tolerate unlawful harassment or discrimination on the basis of sex, gender, race, color, disability, national origin, ancestry, age, citizenship, genetic information or the refusal to submit to a genetic test, past, current, or prospective service in the uniformed services, or any other characteristic protected under applicable federal, Texas, or local law (collectively referred to as Protected Characteristics). These policies also prohibit discrimination and harassment based on pregnancy or related conditions.

If you or someone you know would like help related to an experience involving: 1. Sexual assault, sex-based harassment or discrimination, dating violence, domestic violence, stalking, sexual exploitation, or retaliation for reporting one of these types of prohibited conduct, please visit <www.baylor.edu/titleix>, or contact us at (254) 710-8454, or TitleIX_Coordinator@baylor.edu. 2. Harassment or discrimination (excluding those issues listed in #1) based on Protected Characteristics, please visit <www.baylor.edu/civilrights>, or contact us at (254) 710-8454 or Civil_Rights@baylor.edu.

For pregnancy or related conditions, the University will provide reasonable modifications to policies, practices, or procedures as necessary to ensure equal access to educational programs and activities. Individuals who need to request reasonable modifications should contact the Office of Access and Learning Accommodation. More information on pregnancy and parenting resources can be found here.

The Equity, Civil Rights, and Title IX Office understands the sensitive nature of these situations and can provide information about available on- and off-campus resources, such as counseling and psychological services, medical treatment, academic support, university housing, and other forms of assistance that may be available. Staff members at the office can also explain your rights and procedural options. You will not be required to share your experience. **If you**

or someone you know feels unsafe or may be in imminent danger, please call the Baylor Police Department (254-710-2222) or Waco Police Department (9-1-1) immediately.

Except for Confidential Resources, all University Employees are designated Responsible Employees and thereby mandatory reporters of potential sexual and interpersonal misconduct violations. Confidential Resources who do not have to report include those working in the Counseling Center, Health Center, Athletics Mental Health, and Baylor Psychology Clinic, and the University Chaplain and LHSON Chaplain.