

Syllabus

Welcome!

I love teaching this course, and I am an R enthusiast! At times this course will be challenging, but I promise you the hard work will pay off. Please read this document fully. You may post questions on Piazza or bring them to our first class on 9/2.

1 Course Description

R is the standard tool for performing statistics in biology, health studies, data science, and many other fields. Because of this, **B215 is neither just a biostatistics course nor simply an R programming course**: it is a biostatistics course where R is directly integrated into the learning process. Throughout the course, we use motivating case studies and data analysis problem sets and R labs based on similar challenges to those one finds in scientific research.

Lectures: Focus on statistical concepts and problem solving. We will focus mainly on classic statistics (so-called “frequentist”) and probability theory, but we will also have an introduction to the increasingly relevant Bayesian statistics approach. Topics we will (aim to) cover:

- Sampling & Experimental Design
- Descriptive statistics
- Data visualization
- Probability Theory
- Estimation & Hypothesis Testing
- The Normal Distribution
- Comparing samples & Correlations
- Bootstrapping & Permutation Tests
- Maybe: Frequentist *vs.* Bayesian *vs.* Likelihood approaches

Labs: focus on developing R programming skills through mini lectures and lots of practice. For labs **you will need a laptop with a browser**. Tablets, in my experience, will not work. You can borrow a laptop from the bio department by emailing our admin Andrew (agalleghe[at]brynmaur.edu). Lab topics will include:

- the fundamentals of any programming language, which will be relevant for any future programming languages you decide to learn
- how to run built-in functions in R to run summary statistics, make plots, and perform data analyses in general
- how to use Rmarkdown to make beautiful reports with text and code
- how to write your own functions for custom analyses
- how to troubleshoot/debug your code, arguably the most important skill a programmer must develop

Usually we will have a quick lecture, a guided tutorial, and occasionally a challenge.

2 Textbook

The Analysis of Biological Data (3rd edition), by Whitlock & Schluter. This is a great book with a ton of exercises and online resources (see [R resources](#)). The e-book version is cheaper if you rent it. Collier Library will have one copy on reserve for enrolled students. You may use the 2nd edition but keep in mind that it is your responsibility to determine which chapters/pages/sections correspond to the assigned readings. I may occasionally assign other readings from free sources listed under [R resources](#).

3 Weekly Course Structure

Lectures: Introduce statistical concepts and address questions based on the assigned readings. In-class activities will occur often and count towards participation. Lectures are designed with the assumption that you have completed the pre-assigned readings, enabling us to dive deeper into nuances, applications, and questions.

Labs: R computer lab (bring your laptop) with guided tutorial. R Labs are due by the end of the day.

Problem sets and coding assignments are typically due on Fridays, but always check case by case. Moodle quizzes are typically due weekly on Sundays (soft deadline) though you may retake them later until the hard deadlines (see [Schedule](#)).

Classes will not be recorded. You may not record classes without explicit permission from me.

4 Grade Distribution

Note: The overall course grade visible on Moodle is not a reflection of your current grade for the course. This is confusing for students because the Moodle gradebook requires the instructor to enter percentages for each assignment at the beginning of the course and is somewhat inflexible if changes are required during the semester. Always refer to the list below to understand your current grade.

Component	Weight
Participation	5%
Low-stakes	7%
Problem sets	$5 \times 3 = 15\%$
Coding assignment 1	5%
Coding assignment 2	8%
Midterm 1	30%
Midterm 2	30%

Low stakes assignments. Several assignments in this course will be graded for completion/effort, as they are intended to motivate and aid your learning. These include: DataCamp coding exercises, the weekly R Labs, and multiple choice moodle quizzes based on readings (roughly one quiz per topic). The three lowest grades in the low-stakes assignments category will be dropped. Quizzes will be graded as the average mean of your (unlimited) attempts until the hard deadline and will be open until the end of the course, though you are required to complete the quiz at least once before the soft deadline.

Since problem sets and the first coding assignment are formative¹, i.e, their purpose is learning, I will grade them benevolently following this rubric:

Grade (%)	Description
100	Grades > 90 will be rounded up.
70-90	Complete and various degrees of incorrect answers: 3 + major errors ($\sim 70\%$) to a few minor errors ($\sim 90\%$)
60	Complete but sloppy or excessive amount of major errors.
50	Incomplete (not everything was answered)
0	Plagiarism (including inappropriate LLM use, see Section 9)

¹Formative *vs.* Summative Assessments: Formative assessments are focused on supporting the learning process throughout the course. Summative assessments typically occur at the end of a learning period and assess student mastery of the material and whether they have met learning objectives. Learn more: <https://teachers.institute/assessment-for-learning/formative-vs-summative-evaluation-differences/>

Final grades will likely be assigned as follows (if curving is performed, it could increase but not decrease your grade):

Min %	BMC scale
93	4
90	3.7
87	3.3
83	3.0
80	2.7
77	2.3
70	2.0

5 Problem Sets & Coding Assignments

For Problem Sets and Coding Assignments, you may cooperate with classmates but all work submitted must be your own (see Section 9). They will be due at 11:59 PM on the day denoted on the [Schedule](#) page.

Coding assignment (2 in total) submissions need to be completely reproducible R markdown documents. **If your R markdown file does not compile** it will be considered a late day, and you will be notified and will need to resubmit a R markdown file that does compile/knit. You will be deducted further late days for every day it takes for you to turn in a R markdown file that does knit. Note that the 2 coding assignments are weighed differently.

First coding assignment: You will be allowed to resubmit the first coding assignment (formative) within a week of my handing it back to you if you get less than 90% in your grade. The final grade will be a weighted average of both attempts, with the 2nd attempt having twice the weight of the first attempt. The first coding assignment has lower weight because it is focused on your learning.

Second coding assignment: The 2nd coding assignment will be an assessment of what you have learned in labs and the previous coding assignment. There will be no re-submissions for the 2nd coding assignment.

6 Participation

The course is not the same without you. The best way to participate is to come to class and office hours and ask questions. Shy students can also earn participation points by posting question or answers on Piazza, which can be done anonymously. See @seq-FAQs (Q.2).

7 Late work

Grace days. Each student may use up to 4 “grace days” during the semester to hand in work late, no questions asked and without penalty. To use your grace day(s) for an assignment, email me at least a day before the assignment is due and communicate how many grace days you are requesting. There is no need to provide any explanation as to why you are using your grace days; as long as you have enough days left in the semester to use, and have written to me at least a day in advance, I will approve.

Outside of the grace days, I will accept late work with a reduction in grade as follows: Your maximum score for a late problem set or coding assignment will drop by 10%, every 24h, starting from the deadline. In other words, if you turn in the work 1 minute to 24h late, your maximum score will drop to 90%; if you turn it in 24h and 1 minute to 48h late, it will drop to 80%; and so on. Late quizzes and DataCamp completion will receive a zero but you should still take them for practice/learning/participation purposes.

If you are dealing with an emergency or other issue and believe you may require more leniency, please email me and your dean as soon as possible so that we can make arrangements.

**** Make up policy ****

The course is structured in such a way that you can have some late work without penalty. So there will be no make-up assignments.

See the [Schedule](#) for midterm dates. If you absolutely **must** miss a midterm exam, you will need approval from your dean and myself *ahead of time* (barring last minute severe emergencies) before the exam in order to have a makeup. All make-up exams will be completely different from the in-class ones.

8 Midterms

Both midterms are closed book, no internet. However, you may bring handwritten notes on one standard sheet of paper (front and back), on which you can include whatever you feel might be useful to you..

Some questions in the midterms will be similar to those on problem sets, and there will also be coding questions that you will have to answer by hand. The first midterm will be in class during lab time, and the second midterm will be scheduled during finals week.

Calculator for midterms. You may use a calculator to save you time. A scientific calculator with roots, exponents, etc, is fine, but graphic/function calculators will NOT be accepted. Please plan accordingly. A scientific calculator is ~ \$12 and you might also be able to borrow one from a buddy on Midterm days.

9 Plagiarism

All students enrolled in Bryn Mawr Classes are bound by the The [Bryn Mawr honor code](#). In particular, I call your attention to the sections on plagiarism and AI use in the code, which will be followed here.

BMC honor code on Plagiarism:

In reports and other written work, sources of information and of ideas and opinions not the student's own must be clearly indicated; the source of direct quotations must be acknowledged. Failure to do so constitutes plagiarism.

10 Use of LLMs/GenAI

What the BMC honor code says on LLMs/AI:

Generative AI tools should not be used in the completion of any course assignments, exams, or written work, including laboratory work, unless an instructor for a given course has specifically and explicitly authorized their use for the assignment in question.” (...) “Presenting AI-generated content as original student work, submitting AI-generated content without citing its source in artificial intelligence tools, and relying on generative AI tools for closed-book quizzes, tests, and exams all constitute violations of the Honor Code.

What I say: I personally believe that AI in its current/popular form brings with it serious ethical considerations (see section below). However, I know that most students are likely to reach for AI at least some of the time. For the purposes of this class, the following guidelines apply:

You may use LLMs/GenAI for **coding assignments** ONLY, and only in certain ways:

- a) to brainstorm ideas
- b) to decode cryptic error messages or explain why they might have occurred
- c) to ask specific questions about R-language related things like function parameters, etc.

If you do use LLMs for a coding assignment, you must include citations for each instance where you used it, including the prompt you used (I will provide more specific guidelines with the assignments themselves).

For problem sets and other assignments: You may not use AI, as there is no (ethical) reason to do so. The whole point of these formative assessments is to learn and/or practice for the exams.

10.1 Rationale for the LLM/GenAI policy adopted here

While there are many worthy uses of AI (e.g., predicting protein structures based on amino acid sequences, which has been underway for over a decade), the widespread use of LLMs/GenAI for the most mundane tasks and internet searches has an enormous ecological impact.

Climate Change

Data centers already consume about 2% of the global electricity output (as of December 2024², likely much more now). These data centers emit enormous amounts of CO₂, are mostly powered by non-renewable energy sources – for example, projected growth demands on energy grids to keep up with data center development plans are already delaying the retirement of coal power plants in the U.S.³ – and consume an enormous amount of clean freshwater⁴. The convenience of a quick AI-generated summary costs energy and clean water that seems unjustifiable⁵.

Impacts on critical thinking

Most college students in the U.S. right now use chatGPT/AI regularly. An argument can be made that students need to be trained on proper use of AI. It is almost certain that in your professional life you will use this to some extent (or be pressured to do so for the sake of “efficiency”).

Counterpoints: While constructing effective prompts for AI is indeed a skill to be learned, what is far more difficult is to judge what AI tells you, and that requires knowledge and critical thinking. One of the key purposes of a college education is to develop such critical thinking skills. Research clearly shows that higher confidence in LLMs/GenAI is associated with *less* critical thinking, while higher self-confidence is associated with more critical thinking.^{6 7 8} Additionally: Another purpose of college education is to prepare you for a well-rounded and successful career, and **no one** is looking to hire a person who pastes LLM/GenAI output and presents it as their own work.

²Luccioni, S. (2024, December 18). Generative AI and climate change are on a collision course. Wired. <https://www.wired.com/story/true-cost-generative-ai-data-centers-energy/>

³Garcia, M. (2024). AI Uses How Much Water? Navigating Regulation Of AI Data Centers’ Water Footprint Post-Watershed Loper Bright Decision. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5064473

⁴Kirkpatrick, K. (2023). The carbon footprint of artificial intelligence. Communications of the ACM, 66(8), 17-19. DOI: [10.1145/3603746](https://doi.org/10.1145/3603746)

⁵Luccioni, S., Jernite, Y., & Strubell, E. (2024, June). Power hungry processing: Watts driving the cost of AI deployment?. In Proceedings of the 2024 ACM conference on fairness, accountability, and transparency (pp. 85-99). [10.1145/3630106.36585](https://doi.org/10.1145/3630106.36585)

⁶Lee, HP *et al.* (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In Proceedings of the 2025 CHI conference on human factors in computing systems. <https://doi.org/10.1145/3706598.371377>

⁷Kosmyna, N, *et al.* (2025) Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task.” arXiv preprint arXiv:2506.08872. <https://arxiv.org/abs/2506.08872>

⁸Georgiou, GP. (2025). ChatGPT produces more” lazy” thinkers: Evidence of cognitive engagement decline. arXiv preprint arXiv:2507.00181. <https://arxiv.org/abs/2507.00181>

11 Institutional Support

11.1 Physical and Mental well-being

College is a marathon, not a sprint. You will reap more benefits if you are consistent than cramming an unreasonable amount of work into a short period of time. **If you are struggling**, do not wait: **keep me and your academic dean on the same page** so that we can help you achieve your best. Also, do not hesitate to seek out the services available to you such as the [Bryn Mawr College counseling services](#).

11.2 Title IX

Info from the College [website](#):

BMC strongly encourages all students to report any incidents of sexual misconduct. Bryn Mawr College is committed to providing an inclusive environment, free from sexual and gender-based discrimination. Title IX prohibits discrimination based on sex in any federally funded educational program or activity, and forms the basis for Bryn Mawr's policies and resources regarding sex discrimination. The Bryn Mawr College Sexual Misconduct Policy prohibits Title IX sexual harassment, which includes sexual assault, dating violence, domestic violence, and stalking. Bryn Mawr's Policy is more extensive than Title IX, and also covers other gender-based misconduct. BMC's Policy is more extensive than Title IX, and also covers other gender-based misconduct.

Please be aware that all Bryn Mawr/Haverford employees (other than those designated as confidential resources such as counselors, clergy, and healthcare providers) must report information about such discrimination and harassment to the Bi-Co Title IX Coordinator.

11.3 Students with physical or learning differences

Info from the College website:

BMC is committed to providing equal access to students with a documented disability. Students needing academic accommodations for a disability must first register with Access Services. Students can call 610-526-7516 to make an appointment with the Access Services Director, Deb Alder, or email her at [dalder\[at\]brynmawr.edu](mailto:dalder[at]brynmawr.edu) to begin this confidential process. Once registered, students should schedule an appointment with the professor as early in the semester as possible to share the verification form and make appropriate arrangements.

Please reach out as early as possible (ideally, even before classes start) so that we can make the necessary accommodations. Note that **accommodations are not retroactive and require advance notice to implement**.

Need help with math? The [Bryn Mawr College Q Center](#) supports students who are doing quantitative work in courses across the STEM and Social Science disciplines. The Q Center is a collaborative study space that provides a welcoming location for individual work, study groups, and collaboration with Q Mentors.

11.4 Fostering an Inclusive Learning Space

In an ideal world, it would be possible for scientific practice to be done in a perfectly objective way, free of biases and preconceptions. However, it is not so. Scientists are flawed and human. The history of science is also a history of who got the credit and is dominated by white, male, colonialist, and racist ideas. I will try to highlight less-known historical figures and provide a critical perspective on such matters whenever possible. I see the diversity of backgrounds and identities our students bring as a strength: age, culture, disability, ethnicity, gender, nationality, religion, sexuality, and socioeconomic status. I intend for this course to serve students from diverse backgrounds and perspectives. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally or for other students or student groups. You can do this in person during office hours, by email, or via the mid-semester anonymous survey.

12 FAQs

1. How should I contact the professor?

If you have a general question about course content or assignments, please share it on [Piazza](#). One of your colleagues may know the answer and respond, but our TA and I will also check this regularly. By sharing questions publicly, everyone benefits. Please check Piazza regularly.

For private matters, please email me directly or use the [calendly link](#). I check my email at spaced intervals and will respond to emails within 24 hours Monday through Friday. If you do not hear back within 24 hours (or after the weekend), email again as your email may have been lost. Please plan ahead and do not save any time-sensitive questions for the last minute!

2. What if I have to miss class?

Please let me know you won't be making it to class (lecture or lab) before class starts. Life happens! Everyone is allowed 3 excused absences, no questions asked, no explanations

needed, as long as the student informs me in writing before class. All assignments for class must still be completed on time and with the coordination of relevant parties.

3. What if I cannot turn in an assignment in time?

See Section 7.

4. What if I am requesting an extension because of sickness?

See FAQs 2-3. Most people get sick at least once in a semester. If you have serious health issues that require more leniency, reach out to me and your dean so we can discuss.

5. What if I don't hand in all the low-stakes assignments?

We will have plenty and the three lowest grades will be dropped in this category.

6. Can I use phones/tablets/laptops/etc, in class

Please do not use your phones in class as this hinders instruction and learning both for you and for other students. Tablets with a stylus for notetaking are fine. I would prefer if you only used computers for labs, as the keyboard sounds can be very distracting, but if you feel you must use it for lectures please be mindful of other students.

7. Can I use code from the internet in my assignments?

All work submitted must be your own. Nevertheless, searching the web for help with cryptic error messages is totally normal for any programmer. See “[Getting help with R](#)”. Simply searching “r + how to X” will usually get you somewhere useful. Just cite anything you use to help you solve a problem as described in @ Section 9.

8. Can I use generative AI (e.g., ChatGPT)?

See previous question and Section 9.

9. What if I have accommodations?

Please see Section 11.3 and/or reach out by email if you wish to discuss further.

10. What if I have a conflict with the schedule due to religious holidays?

Check the schedule and if you foresee conflicts due to religious observance, please reach out as soon as possible by email.

11. What if I cannot afford the course materials?

If you have concerns about being able to afford course materials, please reach out by email.

12. I am confused about soft and hard deadlines. For which assignments do they apply?

For certain activities we have suggested deadlines (soft) and hard deadlines. These apply to DataCamp activities and Moodle Quizzes. However, note that for the Moodle quizzes you are still required to take it at least once before the soft deadline as detailed in [Section 4](#).