Curiosity to Question: Research Design, Data Analysis and Visualization 322/382E
Time: Monday and Wednesday. 11:00-12:30
EPS 2.104
Instructor: Dr. Julia Clarke julia_clarke@jsg.utexas.edu (she/her)
Office location: JSG3.216D
Office hours: 1-3 Mondays
Extra hours will be scheduled available via weekly signups and please feel free to email about meeting at other times.
Research Associate Dr. Lucas Legendre lucasjlegendre@gmail.com (he/him) will be giving guest presentations and providing extra meet up times.

Course Description and Learning Objectives: My goal is to empower students in the creative and operational practice of research design and collaborative work in any field in industry or academia. This course is focused on honing principles of experimental design (including formulation of target questions and method choice), scientific writing as well as providing an introduction to tool kits for data analysis, statistical approaches and data visualization. All students in the class will have independent projects that they will be completing throughout the course of the semester. This course has weekly deadlines. We work incrementally, with bite-sized pieces contributing towards completing a project documented in a paper. The final project of the course is a complete, formatted short manuscript for a scientific journal. Graduate students in the class receive additional training in mentorship and are expected to model leadership in the course. All participating students learn the strengths/weaknesses and basics of programming and data visualization in R.

Mentorship and Leadership outcomes: We have monthly community-focused meet-ups during course time to discuss research practice. Course attendance and engagement is essential for these meetups and other course work.

This class carries: Writing and Independent Inquiry flags

Text: None. Course readings will be distributed by date via canvas. A list follows the proposed course schedule- you are expected to consult this document for readings they are not listed as assignments in canvas. However, readings added and changed to suit the progress of the class will be announced via canvas.

Required Equipment: You will need to bring a laptop to each class and have access to Canvas and Google Docs in class. You will need to install R and RStudio installed by Oct. 2.

Please download and install R before you install RStudio:
R: https://cran.r-project.org/
RStudio: https://posit.co/download/rstudio-desktop/
**Date and time of final exam:** There is no final exam. But the final paper is due on the final date/time. **Final Project:** All students will complete a final paper structured as a complete, formatted scientific paper. The paper will be incrementally written throughout the course with feedback provided and will make up 30% of your overall grade. The due date for the final paper is **Dec. 6** at 9am.

**In class peer review:** When this appears on the syllabus you are required to **bring your writing assignment in a sharable format on your laptop.** I recommend printing it before class but this is not required.

**Community.** This design and writing-focused class format succeeds because of the interdisciplinary, supportive community we are able to create. We need to prioritize this. Attendance is expected. Because in-class work is central to the class, lack of attendance will result in a low resulting grade in the course. If circumstances interfere with your ability to be present in course in any way. I will work with you on accommodations. If you must be absent, notify me 1 week in advance.

**Grading:** There will be writing assignments uploaded after peer review that I will provide feedback on (see schedule below and on canvas). These are worth 30% of your overall grade (=30%). Other short assignments for peer review, in class work, R assignments, journal completion, and presentations are worth a total of 30% of your grade. There will be ~3 journal entries across the semester. Class participation in peer review and short meetups with mentors outside of class will count as 10% of your grade. +/- grading. Point values assigned in canvas only approximate this break down.*I do not give any extra credit assignments to individuals, as they are unfair to other students who worked hard on the regular assignments.

**Course plan:** **NOTE reading assignments and topics will shift as the course progress necessitates.** It is your responsibility to follow changes announced to the course plan. These changes will be announced in class or via email so please make sure I have your preferred email contact. Note based on the interest and experience of the class primary and suggested readings may be swapped.

**Classes/topics covered:**

**Week 1.**

**M-Aug. 21.** Building a supportive community for research creativity and success. Empathy, values, your professional goals, and your science. Tools for supporting creativity and project design. Introduction of exercise for Wednesday.

- **Tips: Don't have a project?** What are you curious about? What made you interested in geoscience or your field of inquiry? Brainstorm terms and topics, keywords.
- **Google Scholar** is a good starting point. Read broadly. Look at the datasets and methods of papers close to what you would like to do. Look at how people have approached the questions you are interested in.
- **Set up a meeting with me to go over your project concept or ideas.**

**Think broadly then start narrowing down to one part, region, or one dataset etc.**

- **Tips: Stay organized. Create a project folder on your desktop:**
  - Create a doc. “interesting articles/readings” doc. – with grabs from the cite function in Google Scholar – things you think are relevant. Highlight in yellow those you have not yet read.
  - PDF subfolder you have read and think are relevant [ can have a folder of ones you think may be interesting but not yet read].
  - Working References Cited doc: articles you know you will use in your work that you have read, alphabetical, all formatted the same. APA or Harvard format or that used by your target journal. These can be output from Google Scholar but need to the checked and fixed since there are often errors in the raw output.
  - Notes.doc
  - MainText.doc

**Week. 2. The structure of a scientific paper.**

M-Aug. 28. Dissecting the structure of a scientific paper you like related to your project idea.

In class: “live blogging the structure of a scientific paper” exercise starts.

**Due:** Find a well-written paper in your subject area (using, e.g., Google scholar or UT library), dissect it, and be prepared to explain specifics about what you deduce about the components of a manuscript and what made this particular publication successful or unsuccessful in your view.

Due. Look up the **impact factor** and **Aims and Scope** of the journal in which your article was published. (IF for several (2+) different years – bring to class). In class workshop.

**Week 3. Curiosity to Question: Hypotheses, hypothesis testing, and experimental design.**

**M HOLIDAY**

**W. Sept 6.** Different data types, formulating testable hypotheses. Divergent and convergent thinking exercise. In class work. Honing on your project design.

**Due: One journal entry** on your background in research and scientific writing + what you would like to get out of the class- expectations. Be as specific as possible. Submit via canvas (see specific guidelines on canvas- 1+ pages. Credit/no credit- but a chance for reflection.

*Reading:* find and have read 5 papers related to your research topic. Add these to your working references cited.

*In class:* You will be breaking into groups and sharing your topic and working on the abstract and articulating hypotheses.

**Week 4. Planning your research and refining your question**

**M. Sept 11.** Workshopping your hypothesis. Moving from Big Questions to a specific project.

**Due: create a bibliography (references cited) document** with 5+ papers on your topic and upload it to Canvas (use a consistent format for references – APA format in google Scholar cite. **Due. One line hypothesis statement. Bring to class.**

An example of an online-only citation:


And if there are more than 10 authors you may use "et al."

*Begin writing your abstract (first ½) it is due next class.*


**W Sept. 13. First part of abstract due**— *in-class peer review* (include title, name, institutional affiliation).

*Recommended readings for abstract writing:* Matthews pp. 1-16, Lindsay pp. 49-51 (abstract section).

**Week 5: Research Context: Introduction and Bibliography**

*Discussion, what tests would adequately and exactly address your hypothesis*

**M Sept. 18. Introduction: Context and Justification and Intro to Methods.***

*All Read:* Lindsay pp. 20-27 (“The Introduction”) Matthews Chapter 3, Suggested (pp. 53-75);

*Due: 10-15 papers related to your topic uploaded as fully formatted bibliography*

*Due: Find 1-2 datasets in the supplementary information of 1-2 of these papers. How is the data formatted? What would you do with it to approach your question and test your hypothesis? Be ready to discuss what you found in class*

*Research toolkits and citation workshop—locating the background information you need and synthesizing it in the abstract and introduction. Writing support resources, data repositories, data sharing, reproducibility, and open science.*

*Readings due:*


*Suggested reading: Gopen and Swan, 1990. (in writing docs folder)*

**W. Sept. 20. Break out discussion:** Designing your future and cultures of research practice and mentorship discussion.

*Readings due: All*


And, choose one or skim both: *To start us thinking about lab cultures and mentorship:*


Some starting questions:

Due: reflect on your experiences. What experiences would you describe as mentorship and research mentorship and why? What experience might not be mentorship? What does the culture in the lab or workplace you will lead look like?

Optional: (see mentorship files for other optional articles of interest and links below)


Week 6. The Research Plan, Methods and Data Visualization

M Sept. 25. First draft introduction section due - In class peer review. Use “live blogging” rubric. Make sure all points are hit. Breakout groups to work on hypotheses and planned methods.

W Sept. 27. Revising the Introduction and updating your sources. In class peer review.


Read: Lindsay pp. 28-31 (methods) Introduction to R (on canvas)

Download R and R Studio before class.

R can be downloaded here: https://cran.r-project.org/
and R Studio from here https://www.rstudio.com/

R module 1 will be started in class- Basics of the R environment and prompt, basic math, variables and data types, installing and loading packages, reading in data.

W Oct. 4. How might representing our question via quantitative analysis and data visualization help us think about our questions in new ways?
Due: R module 1 exercise: Basics of the R environment and prompt, basic math, variables and data types, installing and loading packages, reading in data, best practices for setting up a working directory.

Due: Bring data/datasets you are thinking about to class.

Resources (Optional videos and Readings):

Douglas, A., Roos, D., Mancini F., Couto A., Lusseau D. 2023. An Introduction to R. Available at: https://intro2r.com/ (Note: check primarily chapters 1 to 3)

An Introduction to R and RStudio, by Hefin Rhys (Part 1 – you can find Part 2 in the description): https://www.youtube.com/watch?v=IL0s1coNtRk

Week 8. Methods, data collection and formatting and analysis: using R for reproducible research


Due: R module 2: data visualization & communicating results (plotting and data visualization using ggplot, producing publication-ready figures, reporting and reproducible analysis with RMarkdown).

Due: Bring the methods outline, datasets you are working with and data file to class.

Resources (Optional videos and Readings):


W. Oct 11. Data and reproducibility. The right size for your project and aims this semester

**Due: R module 3:** Advanced R (manipulating data with dplyr, writing functions, subsetting, summarizing data, importance of reproducible research); basic statistics in R (t-tests, ANOVAs, assumptions of statistical tests).

**Due:** Bring the methods outline, the datasets you are working with and datasheet to class.

**Resources (Optional videos and References):**


**Week 9 Results and data visualization: Design principles, layout, communicating for differently able audiences, figure preparation, etc.**

M. Oct. 16. Results-communicating in text and images.

**Due- Read:** Lindsay pp. 31-39 (results)
Troubleshooting what you need help with to be ready to represent your findings.

**Due- Read:** *(Note each of these is a single page)*

W Oct. 18. *Figure discussion and workshop. Figure design as a way of honing in on your core message.*

**Due:** Upload a paper with a well-designed figure. Be prepared to explain what it communicates, how it communicates it, and what general philosophies of figure design it touches on *(see readings for previous class).*
Week 10. Results and data visualization: Writing a results section and telling a story with your figures

M. Oct. 23. Peer Review. Draft of written results and one figure or table due, Figure draft discussion and workshop.

Due: Upload Results section and one draft figure and a complete caption.

W. Oct. 25. Communicating your results. In-class workshop: drawing your research project and workshop presentation of results.

Due: Second journal entry due.

Week 11: Mentorship Discussion and Planning for the Discussion and Conclusions


Due: readings and submit discussion points. – see canvas prompts.

W Nov. 1. Discussion and Conclusions

Due: Read: Lindsay pp. 39-49 (Discussion and Conclusions)

Due: Read and dissect a discussion from a paper that you like. Bring that discussion to class.

Week 12 Peer review week

M. Nov. 6. Due Revised Paper Draft!: In class peer review 1 of whole draft (what you have to date)

Upload: what you have. Revise what you can and mock up what you don’t yet.

Note on citing R: These are two separate citations.
You have to cite R, if you use it.


If you want to mention that you used the RStudio environment, you can cite it as well: RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.
Information on citing R modules is listed at the end of the syllabus.

W. Nov. 8. Revision based on feedback from Monday. Peer review 2.

Due-Read Matthews chapter 5 (“Revising structure and style”).

Due- Upload a draft including all parts of the paper.
Revisions of whole paper due post peer review. Graded.

Week 13. Telling Your Story.


Due: Final Journal Entry

W. Nov. 15. Peer review focused on Discussion and Conclusions section/s drafts, fit with abstract.

Due: Draft of discussion and conclusions section for peer review. acknowledgements added.

Week 14: Thanksgiving

Week 15: Presentations of your findings, discussion of implications

M. Nov. 27. 5 slide final presentations! Day 1.
All presentations must be uploaded before class.

W. Nov. 29. 5 slide final presentations! Day 2
All presentations uploaded before class.


W. Dec. 6. Final paper due at 9am.

Other course resources (PDF chapters provided on Canvas):

Writing resources
· Kapitein, Lukas C., et al. (2005). The bipolar mitotic kinesin Eg5 moves on both microtubules that it crosslinks. *Nature* 435: 114-118. (How to construct a nature summary paragraph; [PDF link](#))
· Lindsay, D. (2011). Scientific Writing = Thinking in Words. CSIRO publishing, Australia.

**Analysis and Data Visualization**


**Statistics and R programming**


**Online resources on R (cheat sheets, tutorials)**

· Guidelines to several plot options in R: [https://www.r-graph-gallery.com/](https://www.r-graph-gallery.com/)
• Custom color plates to make your plots look clearer (and prettier): https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html
• Cheat sheet with all basic R functions:
• ...And many other cheat sheets for other packages, R syntax, plot parameters, etc.: https://www.rstudio.com/resources/cheatsheets/
• How to cite R and R packages in your reference list:
  https://www.blopig.com/blog/2013/07/citing-r-packages-in-your-thesispaperassignments/

Mentorship and Future Faculty Skills
• National Academy of Sciences: Adviser, Teacher, Role Model, Friend: On Being a Mentor to Students in Science and Engineering
  http://www.nap.edu/readingroom/books/mentor

CHEATING/UNIVERSITY HONOR CODE: We maintain a zero-tolerance policy on cheating. Scholastic dishonesty includes, but is not limited to, cheating, plagiarism, collusion, using, buying, stealing, and/or divulging the contents of an examination, removing a test from the examination room, substituting for another person, having someone take a test for you, misplacing or damaging property of the University or destroying information so another student may not have materials, falsifying research data, misrepresenting facts including providing false grades or resumes, presenting someone else’s work as one’s own academic work, and falsifying academic records. A full and comprehensive statement about what constitutes academic dishonesty can be found in Appendix C, section 11-802 in the General Information bulletin. The Student Judicial Services office in the Office of the Dean of Students has the responsibility for following up and making the final determination.

You are responsible for knowing and following the UT Policies and the UT Honor Code, available through Canvas when you log in to the course page. The UT Honor Code states:
The core values of The University of Texas at Austin are learning, discovery, freedom, leadership, individual opportunity, and responsibility. Each member of the University is expected to uphold these values through integrity, honesty, trust, fairness, and respect toward peers and community.

POLICY ON Chat GPT: We will be exploring the use of chatGPT in ideation - coming up with a project and look at how it fails at the complex work of scientific research. Please note it especially the latest version) does a terrible job generating accurate text and citations like those required in this class. It often makes up citations that do not exist and delivers data or content that is wholly inaccurate. It is easy to detect. You will not do well in this class if you rely on Chat GPT to write your paper. However, it can be a useful tool at certain stages that we will discuss in class.

EMAIL: All students should become familiar with the University's official e-mail student notification policy. It is the student's responsibility to keep the University informed as to changes in his or her e-mail address. Students are expected to check e-mail on a frequent and regular basis in order to stay current with University-related communications, recognizing that certain communications may be time-critical. It is recommended that e-mail be checked daily. The complete text of this policy and instructions for updating your e-mail address are available at http://www.utexas.edu/its/policies/emailnotify.html.

In this course e-mail will be used as a means of communication with students. You will be responsible for checking your e-mail regularly for class work and announcements. Note: if you are an employee of the University, your e-mail address in Canvas is your employee address.

FOOD AND PHONES: Food is not permitted in lecture or lab. The same rules for drinks that apply in the library apply in the lecture (coffee and other drinks must be in a secure cup or bottle). Cell phones must be off when you are in class or lab. No text messaging in class or lab; if you are messaging etc. on your phone you will be warned once then points deducted from the day's assignment. Laptop computers are required. If you are working on other class assignments or on social networking sites (Instagram etc.) or email during class, you will be warned once then points from the day's assignment deducted.

DISABILITY & ACCESS (D&A)

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think
you may have a disability, and need accommodations please contact Disability & Access (D&A). Please refer to the D&A website for more information: http://diversity.utexas.edu/disability/. If you are already registered with D&A, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

RELIGIOUS HOLY DAYS

By UT Austin policy, you must notify me of your pending absence for a religious holy day as far in advance as possible of the date of observance. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.