INFO 526 Syllabus Fall 2023

Course Description

INFO 526 Data Analysis and Visualization - This course provides an overview of the various concepts and skills required for effective data visualization. It presents principles of graphic design, programming skills, and statistical knowledge required to build compelling visualizations that communicate effectively to target audiences. Visualization skills addressed in this course include choosing appropriate colors, shapes, variable mappings, and interactivity based on principles of color perception, pre-attentive processing, and accessibility.

Course offering

981-2234-1 INFO 526 001 – Data Analysis and Visualization

Instructor Information

Instructor

Dr. Greg Chism,

Assistant Professor of Practice,

School of Information

gchism@arizona.edu

Office: Harvill 420

Office Hours: Mondays, 1-2pm, Harvill 420

Teacher's Assistant

Laura Dozal

PhD Student

School of Information

lwerthmann@arizona.edu

Office: Zoom; available upon request.

Prerequisites

Strong quantitative, statistical, and analytical reasoning abilities are needed to succeed in this course.

This course assumes that this is not your first interaction with working with data in R, using RStudio, and along with version control with Git, and collaboration on GitHub. The course will start with a quick review of the relevant technologies.

Course Format

Live in-person lectures, MW 3:30-4:45pm, Modern Languages, RM 410. Attendance is mandatory.

Course Objective

This course provides an overview of the various concepts and skills required for effective data visualization. It presents principles of graphic design, programming skills, and statistical knowledge required to build compelling visualizations that communicate effectively to target audiences. Visualization skills addressed in this course include choosing appropriate colors, shapes, variable mappings, and interactivity based on principles of color perception, pre-attentive processing, and accessibility.

Learning Outcomes

- Understand the principles of designing and creating effective data visualizations.
- Evaluate, critique, and improve upon one's own and others' data visualizations based on how good a job the visualization does for communicating a message clearly and correctly.
- Post-process and refine plots for effective communication.
- Use visualizations for evaluating statistical models and for statistical inference.
- Master using R and a variety of modern data visualization packages to create data visualizations.
- Work reproducibly individually and collaboratively using Git and GitHub.

Textbooks

Readings for the course will come from the following textbooks. All of them are **freely available online** and you do not need to purchase a physical copy of either book to succeed in this class.

1. [`ggplot2-book`] Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen. [ggplot2: Elegant Graphics for Data Analysis](). (in progress) 3rd edition. Springer, 2022.
2. [`socviz`] Kieran Healy. [Data Visualization: A Practical Introduction](). Princeton University Press, 2018.
3. [`fdv`] Claus O. Wilke. [Fundamentals of Data Visualization](). O'Reilly Media, 2019.
4. [`r4ds`] Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. [R for Data Science](). 2nd edition. O'Reilly, 2022.

Course Schedule

An up-to-date schedule, assignments, and due dates can be found on the course website: [datavizaz.org]().

Course competencies

This course is a core requirement for the MS in Data Science. It will help you master the following competencies:

- MSINFO2: Students will obtain the skills of collecting, manipulating, and analyzing different types of data at different scales, and interpreting the results properly.

- MSINFO3: Students will acquire the skills to communicate the results of their work to interdisciplinary teams, using appropriate visualizations, multi-media, or artistic performance.

Course Community

UArizona Community Standard

All students must adhere to the UArizona Student Rights & Responsibilities: The University of Arizona is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, and accountability. Citizens of this community commit to reflect upon these principles in all academic and non-academic endeavors, and to protect and promote a culture of integrity.

Inclusive community

It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength, and benefit. It is my intent to present materials and activities that are respectful of diversity and in alignment with UArizona's Commitment to Diversity and Inclusion. 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.

Furthermore, I would like to create a learning environment for my students that supports a diversity of thoughts, perspectives and experiences, and honors your identities. To help accomplish this:

- If you have a name that differs from those that appear in your official UArizona records, please let me know! You'll be able to note this in the Getting to know you survey.
- If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with me. If you prefer to speak with someone outside of the course, your academic dean is an excellent resource.
- I (like many people) am still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please let me or a member of the teaching team know.

Communication

All lecture notes, assignment instructions, an up-to-date schedule, and other course materials may be found on the course website: datavizaz.org.

I will regularly send course announcements via email and Slack, make sure to check one or the other of these regularly. If an announcement is sent Monday through Thursday, I will assume that you have read the announcement by the next day. If an announcement is sent on a Friday or over the weekend, I will assume that you have read it by Monday.

Where to get help

- If you have a question during lecture, feel free to ask it! There are likely other students with the same question, so by asking you will create a learning opportunity for everyone.
- The teaching team is here to help you be successful in the course. You are encouraged to attend office hours to ask questions about the course content and assignments. Many questions are most effectively answered as you discuss them with others, so office hours are a valuable resource. Please use them!
- Outside of class and office hours, any general questions about course content or assignments should be posted on the course Slack. There is a chance another student has already asked a similar question, so please check the other posts on Slack before adding a new question. If you know the answer to a question posted on Slack, I encourage you to respond!

Check out the Support page for more resources.

I want to make sure that you learn everything you were hoping to learn from this class. If this requires flexibility, please don't hesitate to ask.

- You *never* owe me personal information about your health (mental or physical) but you're always welcome to talk to me. If I can't help, I likely know someone who can.

- I want you to learn lots of things from this class, but I primarily want you to stay healthy, balanced, and grounded during this crisis.

Lectures

The goal of the lectures is for them to be as interactive as possible. My role as instructor is to introduce you new tools and techniques, but it is up to you to take them and make use of them. A lot of what you do in this course will involve writing code, and coding is a skill that is best learned by doing. Therefore, as much as possible, you will be working on a variety of tasks and activities throughout each lecture and lab. Attendance will not be taken during class but you are expected to attend all lecture and lab sessions and meaningfully contribute to in-class exercises and discussion.

You are expected to bring a laptop to each class so that you can take part in the in-class exercises. Please make sure your laptop is fully charged before you come to class as the number of outlets in the classroom will not be sufficient to accommodate everyone. See the UArizona Libraries loaner technology if you need a loaner laptop.

Assessment

Assessment for the course is comprised of four components: attendance and participation, reading quizzes, homework assignments, and projects.

- **Attendance and participation** is required throughout the semester. Students who attend at least 80% of the lectures and participate regularly during lecture will receive full credit for this portion of their grade. Participation on Slack will also count towards this component.

- **Reading quizzes** (6), due every other week (roughly), completed individually. Each quiz is worth 2% of the grade. Lowest quiz score is dropped.

  Reading quizzes will be linked from the course schedule. They always cover reading that is due since the previous quiz and up to and including the deadline for the given quiz. They're due by 12pm ET (beginning of class) on the indicated day on the course schedule.

- **Homework assignments** (6), due every other week (roughly), completed individually. Each homework assignment is worth 8% of the course grade. Lowest homework assignment score is dropped.

  Homework assignments are due by 12pm ET (beginning of class) on the indicated day on the course schedule.

- **Projects** (2), mid-semester and end of semester, completed in teams.

  - Project 1: Teams will be given a dataset to visualize. Project 1 is worth 15% of the course grade.
  - Project 2: Teams will pick a dataset of interest to them and/or build an R package that implements a new type of data visualization in R. Project 2 is worth 25% of the course grade.

  The deliverables for each project will include a data visualization, a write up of the process and findings, and a presentation. For the second project, you will be encouraged to think beyond a traditional two-dimensional data visualization (e.g., interactive web apps/dashboards, data art, generative art, physical/tangible visualizations, **ggplot2** extensions, etc.).

  Each project will have a peer review component to provide at least one round of feedback during the process of development. Teams will provide periodic peer feedback to their teammates while working on the projects as well as upon completion. The scores from the peer evaluations, along with individual contributions tracked by commits on GitHub, will be used to ensure that each student has contributed to the teamwork.

  All team members must take part in the presentation. Presentations can be given in person in class, or via Zoom if the team prefers. My preference is that the team stick to

one method of delivery (all presenters in person or all presenters on Zoom), but I realize a lot can change throughout this semester, and we'll adjust accordingly.

All work is expected to be submitted by the deadline and there are no make ups for any missed assessments. See [Late work policy] for policies on late work.

Grading

The final course grade will be calculated as follows:

Category	Percentage
Attendance and participation	5%
Reading quizzes	10%
Homework assignments	45%
Project 1	15%
Project 2	25%

The final letter grade will be determined based on the following thresholds:

Letter Grade	Final Course Grade
A	>= 90
B	80 - 89.99
C	70 - 79.99
D	60 - 69.99
E	50-59.99
F	< 60

These are upper bounds for grade cutoffs, depending on the class performance the cutoffs may be lowered but they won't be increased.

Final Projects

Your task for this project is to create *something* related to data visualization.

This is intentionally vague -- part of the challenge is to design a project that showcases best your interests and strengths.

One requirement is that your project should feature some element that you had to learn on your own. This could be a package you use that we didn't teach in class (e.g., a package for

building 3D visualizations) or a workflow (e.g., making a package) or anything else. If you're not sure if your "new" thing counts, just ask!

More information will be provided throughout the semester.

Final Project Date

Tuesday, December 12

Teams

You will be assigned to a different team for each of your two projects. You are encouraged to sit with your teammates in lectures. All team members are expected to contribute equally to the completion of each project and you will be asked to evaluate your team members after each assignment is due. Failure to adequately contribute to an assignment will result in a penalty to your mark relative to the team's overall mark.

You are expected to make use of the provided GitHub repository as their central collaborative platform. Commits to this repository will be used as a metric (one of several) of each team member's relative contribution for each project.

Course policies

Academic honesty

TL;DR: Don't cheat!

Students are expected to abide by The University of Arizona Code of Academic Integrity. 'The guiding principle of academic integrity is that a student's submitted work must be the student's own.' If you have any questions regarding what acceptable practice under this Code is, please ask an Instructor.

Please abide by the following as you work on assignments in this course:

- **Collaboration:** Only work that is clearly assigned as team work should be completed collaboratively.

- The reading quizzes must be completed individually with absolutely no communication with classmates.

- The homework assignments must also be completed individually and you are welcomed to discuss the assignment with classmates at a high level (e.g., discuss what's the best way for approaching a problem, what functions are useful for accomplishing a particular task, etc.). However you may not directly share answers to homework questions (including any code) with anyone other than myself and the teaching assistants.

- For the projects, collaboration within teams is not only allowed, but expected. Communication between teams at a high level is also allowed however you may not share code or components of the project across teams.

- **Sharing and reusing code:** I am well aware that a huge volume of code is available on the web to solve any number of problems. Unless I explicitly tell you not to use something, the course's policy is that you may make use of any online resources (e.g. RStudio Community, StackOverflow) but you must explicitly cite where you obtained any code you directly use (or use as inspiration). Any recycled code that is discovered and is not explicitly cited will be treated as plagiarism. On individual assignments you may not directly share code with another student in this class, and on team assignments you may not directly share code with another team in this class.

- **Generative AI (e.g., ChatGPT):** I am additionally aware of the potential code AI for coding (I taught a workshop on it…). While these tools are amazing, learners should be aware of the impacts that using such tools can have on core competency. David Humphrey, a computer science professor, **wrote about ChatGPT and its potentially negative impacts on core learning.** It is a good read about the pitfalls of using generative AI in an educational context. By using a generative AI, learners may miss the opportunity to discover how something works and why things are done that way. It is also important to note that the iSchool generally bans utilizing ChatGPT and generative AI in our Academic Integrity Policy.

iSchool Academic Integrity Policy

This policy agreed upon by faculty in the UArizona iSchool applies in addition to the Dean of Students' Code of Academic Integrity.

Students in courses at the UArizona iSchool are expected to maintain rigor in their academic performance with intent to learn, practice, and overcome challenges toward personal growth and enrichment. As future professionals in digital environments, iSchool students are also expected to exercise transparency and integrity in collaborations and in the use of tools and resources that may aid completion in assignments for our courses.

Consider the following PROHIBITED practices in this course, unless the instructor has specifically written instructions or permission to do otherwise:

- Posting a question on an online site such as Chegg.com, and copying and pasting some or all of the response into an assessment

- Posting an assessment from the course on online sharing sites such as Course Hero. Aiding other students in violation of academic integrity is also a violation, and is potential copyright infringement.

- Generating and submitting, in whole or in part, text or code through Artificial Intelligence such as ChatGPT, QuillBot, and text summarizers

- Using, in whole or in part, computer code not written by the student (for example, from another student, a book, or the internet) in an assignment or project. This includes using such code in modified or unmodified form.

- Searching for solutions to projects or assignments on the internet or through other tools, when your instructor intended for you to learn the solution through exercises (e.g. Googling for the solution to a question on an assignment).

- Simultaneously submitting the same assignment as another student enrolled into the course without prior permission from the instructor

Exceptions: Clear Instructions will be Provided

In any cases in which this course requires or permits students to use practices in the list above, clear written instructions will specify the tools allowed or required, so students can be certain they are working as instructed. See the UArizona iSchool Academic Integrity Policy, the UArizona Code of Academic Integrity and Syllabus policy for more information.

LLMs and ChatGPT

Large language models (LLMs) like ChatGPT are a type of artificial intelligence (AI) engine that can look like it generates the code you need for R labs and short answer questions. You are encouraged to use ChatGPT to debug code and experiment. However, abuse of ChatGPT can be traced (e.g., failing to give credit or cite ChatGPT when it is used) which could result in your suspension or termination from the course and even your program of study. Keep in mind, too, that while the code may appear legitimate, early studies have shown ChatGPT is not all that accurate with sophisticated coding. Exercise your scholarly discretion and maintain a sense of integrity in your statistical learning journey.

See my policies on this subject above.

Late work & extensions

The due dates for assignments are there to help you keep up with the course material and to ensure the teaching team can provide feedback within a timely manner. We understand that things come up periodically that could make it difficult to submit an assignment by the deadline.

Policy on late work depends on the particular course component:

- Reading quizzes: Late quizzes are not accepted and there are no make ups for missed quizzes.

- Homework assignments: GitHub repositories will be closed to contributions at the deadline. If you need to submit your work late, send a DM on Slack to me to reopen your repository.

  – Late, but same day (before midnight): -10% of available points

  – Late, but next day: -20% of available points

  – Two days late or later: No credit, and we will not provide written feedback

- Projects: The following three components contribute to your project score.

  – Presentation: Late presentations are not accepted and there are no make ups for missed presentations.

  – Write up: GitHub repositories will be closed to contributions at the deadline. If you need to submit your work late, Slack/email me to reopen your repository.

    ∗ Late, but same day (before midnight): -10% of available points

    ∗ Late, but next day: -20% of available points

    ∗ Two days late or later: No credit, and we will not provide written feedback

  – Peer evaluation: Late peer evaluations are not accepted and there are no make ups for missed presentations. If you do not turn in your peer evaluation, you get 0 points for your own peer score as well, regardless of how your teammates have evaluated you.

Waiver for extenuating circumstances

If there are circumstances that prevent you from completing a reading quiz or homework assignment by the stated due date, you may email me (gchism@arizona.edu) before the deadline to waive the late penalty. In your email, you only need to request the waiver; you do not need to provide explanation. This waiver may only be used for once in the semester, so only use it for a truly extenuating circumstance.

If there are circumstances that are having a longer-term impact on your academic performance, please let your academic dean know, as they can be a resource. Please let me know if you need help contacting your academic dean.

Regrade requests

Regrade requests must be made within one week of when the assignment is returned, and must be typed up and submitted via email to me. These will be considered if points were tallied incorrectly or if you feel your answer is correct but it was marked wrong. No regrade will be made to alter the number of points deducted for a mistake. There will be no grade changes after the second project presentations. Note that during the regrade process your score could go up or go down or not change.

No grades will be changed after the project presentations.

"Incomplete" grade

The grade of "I" may be awarded only at the end of a term, when all but a minor portion of the course work has been satisfactorily completed. The grade of I is not to be awarded in place of a failing grade or when the student is expected to repeat the course; in such a case, a grade other than I must be assigned. Students should make arrangements with the instructor to receive an incomplete grade before the end of the term, ..., If the incomplete is not removed by the instructor within one year the I grade will revert to a failing grade.

Tutoring

Tutoring can be found through the UArizona Think Tank.

Attendance policy

Responsibility for class attendance rests with individual students. Since regular and punctual class attendance is expected, students must accept the consequences of failure to attend.

However, there may be many reasons why you cannot be in class on a given day, particularly with possible extra personal and academic stress and health concerns this semester. All course lectures will be recorded and available to enrolled students after class. If you miss a lecture, make sure to watch the recording and review the material before the next class session. Overall this policy is put in place to ensure communication between team members, respect for each others' time, and also to give you a safety net in the case of illness or other reasons that keep you away from attending class.

Note that attendance and participation is part of your grade as well.

Attendance policy related to COVID symptoms, exposure, or infection

Student health, safety, and well-being are the university's top priorities. To help ensure your well-being and the well-being of those around you, please do not come to class if you have symptoms related to COVID-19, have had a known exposure to COVID-19, or have tested positive for COVID-19. If any of these situations apply to you, you must follow university guidance related to the ongoing COVID-19 pandemic and current health and safety protocols. If you are experiencing any COVID-19 symptoms, contact student health at (520) 621-9202. To keep the university community as safe and healthy as possible, you will be expected to follow these and the university guidelines on COVID-19 mitigation. Please reach out to me and your academic dean as soon as possible if you need to quarantine or isolate so that we can discuss arrangements for your continued participation in class.

Notify your instructor(s) if you will be missing up to one week of course meetings and/or assignment deadlines

If you must miss the equivalent of more than one week of class and have an emergency, the Dean of Students is the proper office to contact (DOS-deanofstudents@email.arizona.edu). The Dean of Students considers the following as qualified emergencies: the birth of a child, mental health hospitalization, domestic violence matter, house fire, hospitalization for physical health (concussion/emergency surgery/coma/COVID-19 complications/ICU), death of immediate family, Title IX matters, etc.

Please understand that there is no guarantee of an extension when you are absent from class and/or miss a deadline.

Note: If you've read this far in the syllabus, email me a picture of your pet if you have one or your favorite meme!

Additional university policies

Additional policies can be found at this link (please read through them): https://catalog.arizona.edu/syllabus-policies

Safety on Campus and in the Classroom

For a list of emergency procedures for all types of incidents, please visit the website of the Critical Incident Response Team (CIRT): https://cirt.arizona.edu/case-emergency/overview

Also watch the video available at https://arizona.sabacloud.com/Saba/Web_spf/NA7P1PRD161/common/learningeventdetail/crtfy000000000003560

Important dates

- **Monday, August 21:** Classes begin, Monday schedule
- **Monday, August 28:** Drop/add ends
- **Monday, September 4:** Labor Day, no class
- **Sunday, September 17:** Last day to drop without a W (withdraw)
- **Sunday, October 29**: Last day to withdraw from a class online through UAccess
- **Friday, November 10**: Veteran's Day, limited support available
- **Thursday - Sunday, November 23 - 26**: Thanksgiving recess, no support available
- **Tuesday, December 12**: Project 2 presentations

For more important dates, see the full UArizona Academic Calendar.

Graduate Student Resources

University of Arizona's Basic Needs Resources page for graduate students: http://basicneeds.arizona.edu/index.html