

# Syllabus

## WLF 553 – Reproducible Data Science

### Fall 2024

#### GENERAL INFORMATION

**Schedule:** Tuesday & Thursday 11:00 AM-12:15 PM

**Delivery format:** Hybrid (in person & online synchronous)

**In-person:** Janssen Engineering Building 221

**Online:** <https://uidaho.zoom.us/j/87151755073> (passcode: rds2024)

**Instructor:** Dr. Simona Picardi ([spicardi@uidaho.edu](mailto:spicardi@uidaho.edu), CNR 103F)

**Office hours:** Monday 9:30-10:45 AM, Tuesday & Thursday 1:30-2:20 PM, or by appointment

**Credits:** 3 credits

**Prerequisites:** None

#### COURSE CONTENT

##### Course description

Reproducibility is a pillar of the scientific method. In our day and age, scientists rely on software and code as fundamental tools to carry out their science and ensure it is reproducible. Being proficient in the use of programming tools and effectively apply them to store, process, manage, analyze, and visualize data have become must-have skills to take part in the scientific discourse and to be successful on the job market.

In this course, students will learn best practices in data management and processing for reproducible science. Topics covered will encompass computational tools and techniques to effectively manage data throughout their life cycle, from the moment they get entered into a computer to the moment they are used in a published document. Throughout the course, the students will use a variety of software tools and programming languages to tackle different aspects of data management: spreadsheets for data entry, SQL relational databases for data management, R and the tidyverse for data cleaning, processing, analysis, and visualization, Git for version control of code scripts and data files, and the GitHub platform for code sharing and efficient collaboration.

##### Learning Objectives

The goal of this course is to help students establish their own workflow for reproducible data management that they can apply in their day-to-day research endeavors, by building long-

lasting habits in efficient project organization, data management, and processing. Specific learning objectives include:

- Design functional project structures;
- Organize data in tidy format;
- Design databases that reflect real structure in the data;
- Write code to process data from raw to analysis-ready;
- Maintain code and data integrity with version control;
- Perform automated tasks on data.

### **Prerequisites**

This course has no formal prerequisites. Prior experience with R is beneficial but not required. No experience required for any of the other tools and programming languages.

### **Teaching Philosophy**

One of the major obstacles that students encounter when learning to program is that computational tools are taught incidentally within the context of a course that is centered on a different topic, often statistics. This means students have to split their cognitive load between learning the subject matter and learning the tools to practice it, which is a struggle. This course is designed as an opportunity for students to focus undisturbed on learning the tools. For this reason, it is statistics-free. Another common misconception about learning computational tools is that the best way to learn is to be left alone to confront the code. However, without prior instructions and explanations of how things work, being thrown to the wolves to tackle a programming language is nearly guaranteed to be a frustrating experience. I strive to apply evidence-based pedagogy techniques that are shown to be effective at facilitating learning. These include, among others:

- Setting explicit, measurable goals for each learning activity to facilitate self-assessment;
- Helping students build mental models of each topic before challenging them with practical problem-solving;
- Providing multiple exposure to each topic by combining traditional lectures, live-coding demonstrations, hands-on exercises, and group activities.

### **Instructional Methods**

The course will be delivered in a hybrid format – in person as well as synchronous online. Following a flipped-classroom approach, students will be expected to read the material on each week's topic on their own and then join an in-person class focused on exercises, live coding demonstrations, and group discussions. As a final outcome of the course, students will complete an individual project related to their own research (or any topic of their choice) that implements the techniques learned in class. Weekly lectures are organized chronologically in

the order students are going to need them to work on their project, and each lecture will build off of previous ones.

### **Course Resources**

All software used in this course is open source and freely available online. There is no required textbook. All reading material is provided in the form of a [digital book](#) available on the course website.

### **Code of Conduct**

I am dedicated to providing a welcoming and supportive environment for all people, regardless of background or identity. By participating in this class, participants accept to abide by this Code of Conduct and accept the procedures by which any Code of Conduct incidents are resolved. Any form of behavior to exclude, intimidate, or cause discomfort is a violation of the Code of Conduct. In order to foster a positive and professional learning environment we encourage the following behaviors:

- Use welcoming and inclusive language
- Be respectful of different viewpoints and experiences
- Gracefully accept constructive criticism
- Focus on what is best for the community
- Show courtesy and respect towards other community members

If you believe someone is violating the Code of Conduct, please report it through this [Google Form](#) (reports are anonymous unless you wish to disclose your identity) and I will take the appropriate action to address the situation.

### **Evaluation Methods and Criteria**

The course will include weekly assignments and a final project to be handed in at the end of the semester. Assignments and the final semester project will both count as 50% of the final grade. Many of the weekly assignments will be part of your final project, so working on them each week will make sure you are not left with a pile of work to do at the end of the semester. Because each assignment builds off of previous ones, it's best not to fall too far behind or it will be difficult to catch up. However, if you need more time to complete an assignment, just let me know. You will not need to disclose personal information to get an extension.

### **ASSESSMENT**

Performance will be assessed based on the following criteria:

- Weekly assignments: 50% of grade
- Final project: 50% of grade

Grading Scale: A =  $\geq 90\%$     B =  $\geq 80\%$     C =  $\geq 70\%$     D =  $\geq 60\%$     F =  $< 60\%$

## COURSE POLICIES

**Attendance:** Students are expected to attend class unless otherwise agreed with the instructor. You will simply not be successful in this course if you repeatedly fail to attend class. Absences due to illness or family emergencies will be excused, provided that you let me know in advance when you are not able to attend. Coming to class sick is strongly discouraged: if you are feeling sick, please stay home and communicate with me to discuss alternative ways of keeping up with the course. Students registered for the in-person section of this course are expected to show up to class in person. The online option is reserved for students registered for the online section. However, in-person students can occasionally join online to ensure compatibility with fieldwork schedules, attending while sick, or other emergency situations previously agreed upon with the instructor.

**Staying informed:** It is the student's responsibility to keep up to date with any changes to the course schedule, due dates, etc. Changes will be communicated in three ways: via verbal announcements in class, via announcements on Canvas, and via edits to the Syllabus (posted on Canvas). Students are expected to stay informed by regularly checking these different communication channels.

**Statement on inclusivity:** As a professor at the University of Idaho, I acknowledge the importance of diversity and inclusion and how these attributes contribute to the promotion of a positive educational experience. It is my intent to facilitate a healthy, productive, and safe learning environment where diverse thoughts, perspectives, and experiences are welcomed, and individuals' identities (including, but not limited to: race, sex, class, sexual orientation, gender identity, ability, religious beliefs, etc.) are valued and honored. I recognize that as an educator, it is my responsibility to take the initiative to continually learn about diverse perspectives and identities; therefore, if at any point during the course, you feel uncomfortable or concerned, I am more than willing to discuss suggestions, feedback, and anything else that might improve the general effectiveness of this course.

**Use of AI tools:** There have been amazing developments in AI generative text tools and some are freely available on the web and some come with a cost. These are powerful tools, and it is possible for them to aid in the learning process. My policy is designed to ensure that they aid in the learning process rather than replace the learning process.

I will stipulate how you may use AI tools in the instructions of your assignments and final project, so you are responsible for reading the instructions thoroughly and asking for clarification if necessary. Generally, you are allowed to ask AI to explain pieces of code you don't understand. You are not allowed to use AI to generate new code or assignment solutions. You must download all evidence of your work with the AI tool and append that evidence along with the submission draft of the assignment. Failure to provide evidence, or failure to abide by the limits I have placed on your use of AI tools will result in a failing grade for the assignment

and I must report the incident to the UI Dean of Students as stipulated in Student Code of Conduct (see below).

If I do not stipulate how you may use AI tools, you should assume that use is prohibited.

You must cite your use of AI tools using the following guideline for citation ([The Chicago Manual](#), [APA](#), and [MLA](#) all have recommendations for your adoption).

## UNIVERSITY POLICIES AND RESOURCES

### Student Code of Conduct

#### ARTICLE II- PROSCRIBED CONDUCT.

A. Rules and Regulations. The following list describes actions that detract from the effectiveness of a University community and for which students are subject to disciplinary action. Any student found to have committed or to have attempted to commit the following misconduct is subject to the disciplinary process outlined in FSH 2400:

A-1. Academic Dishonesty. Academic honesty and integrity are core values at a university and the faculty finds that even one incident of academic dishonesty may merit expulsion. Instructors and students are jointly responsible for maintaining academic standards and integrity in university courses. In addition to any disciplinary sanctions imposed under the Code, additional consequences for academic dishonesty may be imposed by the course instructor, including issuing a grade of "F" in the course. Any grade issued by the course instructor, whether as a result of academic dishonesty or not, constitutes an academic evaluation and is not disciplinary action. All instructors must report incidents of academic dishonesty to DOS by email or using the reporting form on DOS website. Acts of academic dishonesty include but are not limited to the following:

a. Cheating includes, but is not limited to, the following:

- (1) using any unauthorized assistance in, or having unauthorized materials while, taking quizzes, tests, examinations or other assignments, including copying from another's quiz, test, examination, or other assignment or allowing another to copy from one's own quiz, test, examination, or other assignment;
- (2) using sources beyond those authorized by the instructor in writing papers, preparing reports, solving problems, or carrying out other assignments;
- (3) acquiring, without permission, tests or other academic material belonging to the instructor or another member of the University faculty or staff;
- (4) engaging in any behavior prohibited by the instructor in the course syllabus or in class discussion; or
- (5) engaging in other behavior that a reasonable person would consider to be cheating.

b. Plagiarism includes, but is not limited to, the following:

- (1) using, by paraphrase or direct quotation, the published or unpublished work of another person without full and clear acknowledgment;
- (2) using materials prepared by another person or agency engaged in the selling of term papers or other academic materials without prior authorization by the instructor; or
- (3) engaging in other behavior that a reasonable person would consider plagiarism.

### **Department of Fish & Wildlife Sciences Policy on Plagiarism**

“A fundamental goal of education is to produce students who can evaluate ideas – both analysis and synthesis – and who can produce significant original thoughts. Plagiarism is simply repeating words or thoughts of other people, without adding anything new. Therefore, a student who submits plagiarized text – in addition to the wrongful conduct – does not demonstrate the level of understanding and skill that an educated person is reasonably expected to have.” (R.B. Standler. 2012. Plagiarism in colleges in USA, <http://www.rbs2.com/plag.htm>).

“Plagiarism means using another’s work without giving credit. If you use others’ words, you must put them in quotation marks and cite your source. You must also give citations when using others’ ideas, even if you have paraphrased those ideas in your own words.” (<http://sja.ucdavis.edu/files/plagiarism.pdf>)

“Many people think of plagiarism as the simple copying of someone else’s words with the intent of passing them off as one’s own. While that is certainly a type of plagiarism, it is by no means the definition of it. Plagiarism is a kind of intellectual theft, but rather than being a black-and-white issue, plagiarism comes in various shades of gray. Instances of plagiarism can range from the accidental misuse of a punctuation mark to the willful purchasing of a prepared essay.” (Wake Forest University web-site and tutorial on how to avoid plagiarism <http://zsr.wfu.edu/instruction/recognizing-and-avoiding-plagiarism/>)

“Several authors of law review articles agree that intent is irrelevant in determining whether plagiarism occurred. It is no defense for the plagiarist to say “I forgot.” or “It is only a rough draft.”” (R.B. Standler. 2012. Plagiarism in colleges in USA. <http://www.rbs2.com/plag.htm>)

Some potentially useful web sites (in addition to those cited above):

- <https://owl.english.purdue.edu/owl/resource/747/08/> Purdue University Online Writing Lab – how to cite electronic sources
- <http://www.uidaho.edu/class/english/plagiarism/policies> UI English Department site on plagiarism
- [http://writing.wisc.edu/Handbook/QPA\\_plagiarism.html](http://writing.wisc.edu/Handbook/QPA_plagiarism.html) Univ. of Wisconsin The Writer’s Handbook – Avoiding Plagiarism

- <https://www.scribbr.com/plagiarism/types-of-plagiarism/> Examples of what is and is not plagiarism
- <http://www.plagiariized.com> Commercial site for detecting plagiarism

Plagiarism is a serious issue. Do your best to avoid it. If you are uncertain about how to cite sources, or have other questions about potential cases of plagiarism, visit with your course instructor prior to handing in an assignment.

**Counseling and Testing Center:** The Counseling and Testing Center ([www.uidaho.edu/current-students/ctc](http://www.uidaho.edu/current-students/ctc)) offers students access to a wide range of counseling services, resources, and referrals, such as testing services, outreach and consultation, and psychiatric services. Contact the Center by calling 208-885-6716 or emailing [ctc@uidaho.edu](mailto:ctc@uidaho.edu).

**Center for Disability Access and Resources (CDAR):** The University of Idaho is committed to ensuring an accessible learning environment where course or instructional content are usable by all students and faculty. If you believe that you require disability-related academic adjustments for this class (including pregnancy-related disabilities), please contact Center for Disability Access and Resources (CDAR) to discuss eligibility. A current accommodation letter from CDAR is required before any modifications, above and beyond what is otherwise available for all other students in this class will be provided. Please be advised that disability-related academic adjustments are not retroactive. CDAR is located at the Bruce Pitman Building, Suite 127. Phone is 208-885-6307 and e-mail is [cdar@uidaho.edu](mailto:cdar@uidaho.edu). For a complete listing of services and current business hours visit <https://www.uidaho.edu/current-students/cdar>.

**Healthy Vandals policy:** Please visit the [University of Idaho COVID-19 webpage](#) often for the most up-to-date information about the UofI's response to Covid-19.

**Vandal Food Pantry:** The [Vandal Food Pantry](#) is a free resource stocked weekly with food, grocery bags, and various hygiene items. Its eight locations across campus are accessible during building hours and open to all. Please take what you need.

**Green Dot safety program:** It's up to all of us to make a safer campus. Vandal Green Dot is a program that helps students learn about the power of the bystander, how to recognize potentially risky situations, and realistic ways to intervene. Together we can bring down the number of people being hurt by interpersonal violence on our campus. No one has to do everything, but everyone has to do something! Learn more and get involved by visiting [UI's Green Dot Safety Program](#) or emailing [green@uidaho.edu](mailto:green@uidaho.edu).

**Firearms:** The University of Idaho bans firearms from its property with only limited exceptions. One exception applies to persons who hold a valid Idaho enhanced concealed carry license, provided those firearms remain concealed at all times. If an enhanced concealed carry license holder's firearm is displayed, other than in necessary self-defense, it is a violation of University policy. Please contact local law enforcement (call 911) to report firearms on University property. University of Idaho leadership remains committed to maintaining a safe work, living, and learning environment on campus. We will not tolerate any threatening use of firearms or any other weapons. While authorized license holders may have familiarity and be at ease

carrying a loaded firearm, we ask that they be aware that many people are not familiar with handguns and are uncomfortable in their presence.

**Library:** The library website has many databases that will help you find relevant and reliable books, articles, images, and more. Don't hesitate to contact a librarian for research assistance.

- [UIDAHO Library](#)
- [Help - Reference Services](#)
- [Help for Distance Ed Students](#)

**Technology Help:** The UI Help Desk provides many technology related services to UI students:

- Phone: (208) 885-HELP (208-885-4357)
- Email: [helpdesk@uidaho.edu](mailto:helpdesk@uidaho.edu)
- Website: [www.uidaho.edu/infrastructure/its](http://www.uidaho.edu/infrastructure/its)

**University of Idaho Land Acknowledgement:** The University of Idaho Moscow campus is located on the homelands of the Nimiipuu (Nez Perce), Palus (Palouse) and Schitsu'umsh (Coeur d'Alene) tribes. We extend gratitude to the indigenous people that call this place home, since time immemorial. The University of Idaho recognizes that it is our academic responsibility to build relationships with the indigenous people to ensure integrity of tribal voices.

## SCHEDULE

\*Assignments marked with an asterisk pertain to the final semester project.

Week	Date	Topic	Readings	Assignments
1	Tuesday, August 20, 2024	Intro to Reproducible Data Science		
1	Thursday, August 22, 2024	Project organization	Chapter 1 (1.1 through 1.4)	Install software, choose final project topic*
2	Tuesday, August 27, 2024	Paths and file names	Chapter 1 (1.4 through 1.7)	
2	Thursday, August 29, 2024	Version control (Git)	Chapter 2	Set up final project directory structure and initialize git repository*
3	Tuesday, September 3, 2024	Version control (Git), continued		



3	Thursday, September 5, 2024	Version control (GitHub)	Chapter 3	Set up GitHub repository for final project*
4	Tuesday, September 10, 2024	Version control (GitHub), continued		
4	Thursday, September 12, 2024	Tidy data and best practices in the use of spreadsheets	Chapter 4	Design tidy spreadsheet for assigned case study
5	Tuesday, September 17, 2024	Database structures and design	Chapter 5	
5	Thursday, September 19, 2024	Database structures and design, continued		Design final project database structure*
6	Tuesday, September 24, 2024	SQL queries	Chapter 6 (6.1 through 6.2)	
6	Thursday, September 26, 2024	SQL queries		Solve assigned SQL queries
7	Tuesday, October 1, 2024	Database building	Chapter 6 (6.3)	
7	Thursday, October 3, 2024	Database building, continued		Build final project database*
8	Tuesday, October 8, 2024	Interfacing databases and R with RSQLite	Chapter 7	
8	Thursday, October 10, 2024	Dynamic documents with Rmarkdown	Chapter 8	Write description of final project in Rmarkdown*
9	Tuesday, October 15, 2024	E-books with bookdown	Chapter 9 (9.1)	
9	Thursday, October 17, 2024	Automatically generated websites with GitHub Pages	Chapter 9 (9.2 through 9.3)	Publish final project website*
10	Tuesday, October 22, 2024	NO CLASS		

10	Thursday, October 24, 2024	Foundations of R programming	Chapter 10 (10.1 through 10.2.14)	Solve assigned R tasks
11	Tuesday, October 29, 2024	Automating repeated tasks in R	Chapter 10 (10.2.15 through 10.2.16)	
11	Thursday, October 31, 2024	Troubleshooting in R	Chapter 11	Fix my code
12	Tuesday, November 5, 2024	Data wrangling with tidyverse	Chapter 13	
12	Thursday, November 7, 2024	Data wrangling with tidyverse, continued		Solve assigned tidyverse tasks
13	Tuesday, November 12, 2024	Data wrangling with tidyverse, continued		
13	Thursday, November 14, 2024	Data wrangling with tidyverse, continued		Solve assigned tidyverse tasks
14	Tuesday, November 19, 2024	Data visualization with ggplot2	Chapter 14	
14	Thursday, November 21, 2024	Data visualization with ggplot2, continued		Make assigned visualizations
15	Tuesday, November 26, 2024	Dates and times in R	Chapter 15	
15	Thursday, November 28, 2024	THANKSGIVING		
16	Tuesday, December 3, 2024	Geospatial data management in R	Chapter 16	
16	Thursday, December 5, 2024	Final project case studies + AMA		
17	Friday, December 13, 2024	NO CLASS		Final project due*