

STAT 498C: Agile Data Science

Winter Quarter, 2025

Instructor: Duncan Wadsworth, PhD (he/him), wduncanw@uw.edu

Office Hours: After class or by appointment via Zoom

Course Description

Agile Data Science adapts the tools and processes of Agile Software Development. Students will learn those tools and processes in a project-based environment which will reward collaboration in small groups working toward a final data science deliverable. The project may be an analysis, a dashboard, or a competitive predictive model; suggested projects will be provided but students may propose their own. Class sessions will be a mix of lectures and active group work. The goal of the course is for students to become skilled in the collaborative and iterative nature of data science, emphasizing that these skills are best developed through hands-on practice.

Key learning outcomes

1. Students will exercise the process of data science, including:
 - a. documented planning
 - b. incremental execution
 - c. mid-process communication
 - d. summary presentation
2. Students will learn tools which facilitate collaboration and communication, including:
 - a. version control (git & GitHub)
 - b. project tracking (Kanban project boards, itemized)
 - c. documentation (word processing and presentation, emphasis on text-based formats)
 - d. cloud compute, i.e. using someone else's computer for storage and computations

Ten-week class schedule

1. Introduction to Agile Data Science
2. Data Science Process and Project Planning
3. Project Management
4. Agile Data Science and git Intro
5. GitHub and git Best Practices
6. Mid-term Project Check-in
7. Organization and Feedback
8. Communicating Data Science
9. Project and Story Time
10. Final Presentations
11. Finals Week

Communication Expectations

Email will be the primary form of communication. However, later in the course we will utilize GitHub Issues and Pull Requests for communication. Additionally, we may use in-document comments; to ensure delivery and attention, always tag, e.g. @wduncanw, the intended recipient in your comments.

Instructor will try to respond to email within 48 hours. If a response is not received within 48 hours, please send a reminder. When necessary, sparingly use an [URGENT] tag in the subject line

Class resources

Course resources are documented freely online, including but not limited to:

- Agile Software Development Wikipedia page: https://en.wikipedia.org/wiki/Agile_software_development
- Online git documentation: sections of <https://git-scm.com/book/en/v2> and <https://git-scm.com/docs>
- Online GitHub documentation: sections of <https://docs.github.com/en>
- The “Happy Git With R” website: <https://happygitwithr.com/>
- Jenny Bryan’s Stat 545 course at UBC: <https://stat545.com>
- Hadley Wickham’s Stat 405 course at Rice: <http://stat405.had.co.nz>
- Data Carpentry Workshop materials: <https://datacarpentry.org/r-socialsci/>

Generative AI Policy

Students are encouraged to use generative AI tools in their projects. However, students are expected to be aware of the potential for misuse of these tools and to use them responsibly. Furthermore, if students choose to use such tools, they are *strongly encouraged* to keep logs of the prompts and the output and be prepared to discuss both in class. The logs of these prompts and a summary of their output should be kept in a file named `generative-ai-logs.md` in the root of the project repository. An award will be given to the two teams with the best logs. Best defined as most comprehensive and/or most entertaining.

An example of what should be in the `generative-ai-logs.md` file

While I was writing slide templates for this course, I used a generative AI tool to generate text for the slides.

```
2024-12-04 19:23:33, GitHub Copilot
```

```
PROMPT[Before we get started, I must make clear this material is not an endorsement of any particular product or service. The contents of this lecture are sourced from my own study and experience and separate from my primary employer. All opinions are my own.]
```

```
COMPLETION[I am not a financial advisor. I am not a lawyer. I am not a medical doctor. I am not a professional athlete. I am a student. I am a
```

teacher. I am a learner. I am a human being. I am a human doing. I am a human becoming. I am a human. I am. I.]

Evaluation criteria

An emphasis will be placed on planning, effort, and completion.

	Percent of Final Grade	When
Team Scrums and Labs (equally weighted)	20%	14x During Quarter
Homework (equally weighted)	20%	7x During Quarter
Mid-point Check-in	20%	Week 6
Final Project & Presentation	30%	Week 10
Peer Feedback	10%	Finals Week

Team Scrums and Labs: Students will be expected to participate in scrums and labs during each class session. Labs and scrums will be graded on a pass/fail basis. Students will be expected to complete the lab exercises according to the exit criteria and to submit them via 1:1 document with instructor. Students who miss a lab or scrum will be expected to make it up by the end of the quarter.

Homework: Students will be expected to complete homework assignments on time. Late homework will be accepted with a 10% penalty per day late. Homework will be assigned on Thursday and due the following Thursday, by 11:59pm.

Mid-point Check-in: Students will be expected to present their project plans and progress to the class. Presentations will be graded on a needs work/acceptable/excellent basis. Students who miss the check-in will be required to present during office hours.

Final Project & Presentation: Students will be expected to complete a final project and presentation. The project will be presented with teammates and summarize the work the team has done throughout the quarter. It must showcase the team's planning, effort, and completion.

Peer Feedback: Students will be expected to provide feedback to their teammates. Feedback will be graded on a pass/fail basis. Students who do not provide thorough feedback will receive a failing grade for this portion of the course. Students with substantial negative feedback will be given the opportunity to provide additional context but may receive a failing grade for this portion of the course.

Accommodations

Inclusivity

Among the core values of the university are inclusivity and diversity. Inclusivity in our context means that assignments and activities should be accessible to all students, regardless of race, gender, income, ability, beliefs, and other ways that people distinguish themselves and others. I am committed to this value; If you have any concerns about the inclusivity of any aspect of the course, please contact me so we can discuss it.

Religious Accommodations

Washington state law requires that UW develop a policy for accommodation of student absences or significant hardship due to reasons of faith or conscience, or for organized religious activities. The UW's policy, including more information about how to request an accommodation, is available at [Religious Accommodations Policy](#). Accommodations must be requested within the first two weeks of this course using the [Religious Accommodations Request form](#).

Disability Accommodations

Your experience in this class is important to me. It is the policy and practice of the University of Washington to create inclusive and accessible learning environments consistent with federal and state law. If you have already established accommodations with Disability Resources for Students (DRS), please activate your accommodations via myDRS so we can discuss how they will be implemented in this course.

If you have not yet established services through DRS, but have a temporary health condition or permanent disability that requires accommodations (conditions include but not limited to; mental health, attention-related, learning, vision, hearing, physical or health impacts), contact DRS directly to set up an Access Plan. DRS facilitates the interactive process that establishes reasonable accommodations. Contact DRS at disability.uw.edu.