DATA 510
Capstone

Tues, 200 Market, Portland, OR
Summer 2024

Jed Rembold, Ph.D.
jjrembold@willamette.edu
http://willamette.edu/~jjrembold/classes/data510
Ford 214
Office Hours: By appointment virtually or just catch me online
Phone: (503) 370-6860

This syllabus is very subject to change or adaptation as the semester progresses!

Course Description: Over the course of the semester, you will propose, plan, and execute an actual data science project. The project provides an opportunity to integrate all of the core skills learned throughout the program, and to develop a portfolio piece that can help with your career aspirations. Projects must be consequential in nature–i.e., have a real (or potential) impact on some organization or the world. Grades will be based on milestone assignments and a panel of judges from industry.
Classes will be a mix of professional development activities, lecture and instruction, guest speakers, and group work and mentoring.
Prerequisite(s): DATA 501, DATA 502, DATA 503, DATA 504
Note: A minimum grade of C- is required for this course to count toward university credit.
Credits: 4.0

Text: There is no official text this semester. Handouts will be distributed as necessary.

Course Objectives:
Over the semester, students will gain working knowledge in:

1. Implementing state-of-the-art software development practices to support publication and presentation of a data science project.
2. Intelligently discussing the latest developments in data science and their potential impact on organizations, the economy, and the world.
3. Proposing and defending a data science-based solution to a problem of consequence for an organization, the economy, or the world
4. Communicating the implications of one's findings in a clear and non-technical manner–both in person and in writing.
5. Demonstrate a clear understanding of data science topics and techniques in a professional interview setting.

Grade Weighting

Professional Development Activities	20%
GitHub Pages Portfolio	5%
Project Proposal	10%
Data Summary	10%
Rough Draft Write-up	10%
Final Presentation	20%
Final Draft Write-up	20%
Weekly Check-ins	5%

Letter Grade Distribution:

>= 92.00	A	72.00 - 77.99	C
90.00 - 91.99	A-	70.00 - 71.99	C-
88.00 - 89.99	B+	68.00 - 69.99	D+
82.00 - 87.99	B	62.00 - 67.99	D
80.00 - 81.99	B-	60.00 - 61.99	D-
78.00 - 79.99	C+	<= 59.99	F

Student Learning Objectives (SLO):
Upon completion of the course, students should be able to:

- Navigate the job market and leverage their unique skills and talents to excel in professional interview situations. Demonstrated through completion of the professional development activities.
- Advertise themselves to potential employers, showcasing their strengths and skills in a polished and professional manner. Demonstrated through professional development activities, the GitHub pages profile, and the final write-up.
- Plan out from start to finish the major pieces of a large-scale data science research project, including estimated timelines. Demonstrated through the project proposal and later milestone deadlines.
- Execute a scheduled research plan, meeting deadlines and showcase a polished project at the end. Demonstrated through each of the milestone deadlines.
- Present research findings and the research process to technical and non-technical individuals in an interactive, verbal and visual way. Demonstrated through the professional development activities and the final presentation.
- Write up research findings and the research process in a clear and concise manner suitable for modern publication. Demonstrated through the rough and final draft write-ups.
- Work effectively with others to split workloads to be able to meet deadlines. Demonstrated through project milestones and the final presentation.

Course Assessment:

- **Professional Development Activities**
  - Every other week the first portion of class is set aside for professional development. Ruthie will be having students would through a variety of tasks and situations to help ensure that students are as prepared as possible for entering the job market. Even if students are not actively looking for a new job, the tech-world is a swiftly changing place, and most tech-workers switch jobs multiple times over their career. So all will be expected to participate.
- **GitHub Pages Portfolio**
  - Modern tech-workers and job seekers need a public facing portal in which they can showcase their skills and experience. While there are many ways to go about this, one of the simplest is by setting up a portfolio using GitHub pages. Doing so will be introduced and covered in class, as well as some futher actions students colud take to more extensively customize their portfolio.
- **Project Milestones**
  To keep projects on pace, and to further mimic the more consistent deadlines that are commonly found in the workplace, there are several milestone laid out for the projects over the course of the summer:
  - *Project Proposal* - The project proposal is a short write-up that details all of the planned ideas and steps to bring a project to fruition. That include the research question itself, data sources to support that question, and planned statistical and machine learning analysis. While projects are free to pivot as necessary, the proposal should serve as the default fall-back and act as a guiding light throughout the summer.
  - *Data Summary* - The data summary demonstrates that all the data necessary to answer the research question has been gathered and organized, or pipelines have been implemented that are actively gathering and organizing the data without further interaction necessary. Past this milestone, no further active work should need to be done to bring in new data.
  - *Rough Draft Write-up* - The rough draft write-up represents the first complete draft describing the project's question, methodology, and findings. It will be written in a fashion condusive to online publishing, and should include sections on all of the project requirements.
- **Project Deliverables**
  At the end of the semester, two major things are required:
  - *Final Presentation* - A final 15-minute professional presentation will be given to peers, faculty, and representatives from industry. This is a chance to highlight the work done over the summer and to discuss a project's findings. Final presentations will be joint across both cohorts and happen on the last evening of classes.
  - *Final Write-up* - A final draft of the write-up represents the concluding version of a projects question, methodology, findings and future. It should be extensively proofread and polished, as it is the permanent record of the work the students have accomplished. It will be published online and linked to student portfolios.

- **Check-ins**
  - The summer is a long time and it can be easy for things not to get done on long term projects. Actively thinking about such things tends to help, and thus there will be weekly online check-in forms that students will fill out. These questionnaires will touch on what has been accomplished in the previous week, how student's are feeling about the progress that has been made, and plans for the upcoming week. These are scored purely on a credit/no-credit metric, and will only be up and available over the weekends.

Course Policies:

Late Work Policy

I understand that sometimes things come up, and you are unable to meet a deadline. However, deadlines commonly exist to ensure that you are in a position to achieve success in the future, and workplaces often do not have the flexibility that academics have in allowing extensions. In an effort to maintain some flexibility while still having firm deadline, you have a 24 grace period past each deadline in which you can submit without penalty. Past that you lose 20% of that deliverable's worth each day. Final presentations can not utilize this 24-hour grace window, as we have a single class slot set aside for presentations. The weekend check-ins are also exempt from this policy.

Incomplete Policy

An incomplete grade will only be granted in the case of prolonged illness or family emergencies that remove the student from the campus for an extended time period during the latter portion of the semester. Under no situations will an incomplete be granted due to a student falling behind through lack of motivation, understanding, or time management skills. If you are concerned about your progress and how you are doing in the class, please come visit me! We can sort out where you are struggling and work out a plan to get you back on track.

Classroom Conduct

As an educational institution, Willamette is committed to support the ideals and standards that help create a constructive and healthy learning community. That requires, among other things, encouraging positive classroom behaviors, discouraging disruptive classroom behaviors, and setting clear standards for both of those things.

To that end, constructive classroom behaviors are those that support learners and teachers in an environment that promotes trust, respect, and collaborative learning.

Disruptive classroom behaviors are those that undermine or interfere with the abilities to learn and teach. Clear examples of disruptive behaviors include, but are not limited to:

- Interrupting others or persistently speaking out of turn
- Distracting the class from the subject-matter or discussion at hand
- Making unauthorized recordings or photos of a class meeting or discussion (except as permitted as part of an Accessible Education Services-mandated accommodation)
- Any physical threat, physical, psychological, or sexual harassment, ridicule, or abusive act towards a student, staff member, or instructor in a classroom or related setting.

Willamette Policies:

Academic Honesty

Cheating is defined as any form of intellectual dishonesty or misrepresentation of one's knowledge. Plagiarism, a form of cheating, consists of intentionally or unintentionally representing someone else's work as one's own. Integrity is of prime importance in a college setting, and thus cheating, plagiarism, theft, or assisting another to perform any of the previously listed acts is strictly prohibited. I may impose penalties for plagiarism or cheating ranging from a grade reduction on an assignment or exam to failing the course. An instructor can also involve the Office of the Dean of the College of Liberal Arts for further action. For further information, visit: http://www.willamette.edu/cla/catalog/resources/policies/plagiarism_cheating.php.

This can be particularly problematic in programming courses, so know that I will be keeping an eye out for it. Do your own work, and always indicate if you have worked with someone else.

Time Commitments

Willamette's Credit Hour Policy holds that for every hour of class time there is an expectation of 2-3 hours work outside of class. Thus, for a class meeting three hours a week you should anticipate spending 6-9 hours outside of class engaged in course-related activities. Examples include study time, reading and homework, assignments, research projects, and group work.

Diversity and Disability

Willamette University values diversity and inclusion; we are committed to a climate of mutual respect and full participation. Our goal is to create learning environments that are usable, equitable, inclusive and welcoming. If there are aspects of the instruction or design of this course that result in barriers to your inclusion or accurate assessment or achievement, please notify me as soon as possible. Students with disabilities are also encouraged to contact the Accessible Education Services office in Matthews 103 at 503-370-6737 or accessible-info@willamette.edu to discuss a range of options to removing barriers in the course, including accommodations.

Tentative Course Outline:

The weekly coverage might change as it depends on the progress of the class. However, I highly recommend you follow along with the reading, as it makes a large difference!

Week	Date	Chapter	Description	Due
1	Tue, May 14		Syllabus and Research vs EDA Project Brainstorming	
2	Tue, May 21		Prof Dev: Targeting and Networking Partner/Group Meet-ups and Brainstorming	
3	Tue, May 28		TBD Work on Proposals	
4	Mon, Jun 03 Tue, Jun 04		Prof Dev: Behavioral Interviews Guest Faculty - Data Acquisition work	Project Proposal
5	Tue, Jun 11		Prof Dev: Internal Mock Interviews Scraping and modeling refresher	
6	Tue, Jun 18		Publishing a homepage/portfolio with GH Pages TBD	
7	Tue, Jun 25		Prof Dev: Prepping for technical interviews Data Summary work	
8	Mon, Jul 01 Tue, Jul 02		Technical writing with Quarto Statistics Consultations	Data Summary
9	Tue, Jul 09		TBD Machine Learning Consultations	
10	Mon, Jul 15 Tue, Jul 16		**External Mock Interviews**	Personal Portfolio
11	Tue, Jul 23		Prof Dev: Alumni Panel Alumni Consultations	
12	Mon, Jul 29 Tue, Jul 30		Prof Dev Drop-ins / TBD Data Visualization Consulting	Rough Draft Write-up
13	Tue, Aug 06		Prof Dev Drop-ins / TBD Peer Presentation Practice	
14	Tue, Aug 13 Thu, Aug 15		**Project Presentations**	 Final Draft Write-up