

DATS 4001 – Data Science Capstone



COURSE AND CONTACT INFORMATION

Term	Spring 2026	
Section	11 (CRN: 47830)	10 (CRN: 46052)
Time	Mon (3:30–6:00 p.m.)	Mon (6:10–8:40 p.m.)
Location	PHIL 110	PHIL 110

INSTRUCTOR

Name: Dr. Sushovan Majhi

Office: Samson Hall 313 (easier to enter from Corcoran 2nd floor, through the end-of-hallway stairs)

Email: s.majhi@gwu.edu / Phone: 202-994-1235

Office hours: Mondays 2–3 pm

COURSE DESCRIPTION

The goal of the Capstone Project is for the students to apply the knowledge acquired throughout the GWU Data Science undergraduate program to solve a real-world problem, using real data. It typically starts with developing S.M.A.R.T. questions and goes through the entire data science life cycle in multiple iterations. The students will work individually, with topics that interest them. The project topics need final approval from the instructor.

LEARNING OUTCOMES

As a result of completing this course, students will be able to:

1. understand and execute complete data science life cycles in solving real-world problems to provide findings and solutions;
2. work with industry, government, non-governmental organizations (NGOs), or academic research groups to offer analyses and solutions to their problems;
3. interpret and communicate their work in organized and efficient manners to a variety of audiences orally;
4. interpret and communicate their work in a well-organized written format, in the form of a report, journal paper, or conference paper;
5. respond to questions and critiques in poster and oral presentations;

COURSE PREREQUISITES

DATS 1001, DATS 2101, DATS 2102, DATS 2103, and DATS 2104.

PROJECT REQUIREMENTS

1. Demonstrate ability to carry out a data science project from end to end (life cycle);
2. Demonstrate proficiency in preparing and executing a presentation.
3. Demonstrate the ability to carry out literature research and summarize the state of the art.
4. Demonstrate the ability to translate project objectives into a realistic work plan.
5. Demonstrate ability to design and implement required software using tools such as Python, R, SQL, as well as traditional programming languages such as C, C++, and Java;

COURSE SCHEDULE

Date	Topics	Assignment(s) Due
Jan 12	Capstone overview, Getting to each other Capstone GitHub Repo, Data Science Lifecycle Ideas for a proposal	Project Proposal (draft)
Jan 19	Martin Luther King Day (no class) [Receive comments on Proposal]	
Jan 26	Elevator pitch - Peer feedback	Project Proposal Final
Feb 2	Weekly standup #1 Data collection review	Weekly update #1
Feb 9	Weekly standup #2 EDA review	Weekly update #2
Feb 16	President's Day (no class)	
Feb 23	Weekly standup #3 Data visualization review	Weekly update #3
Mar 2	Weekly standup #4 Modeling review	Weekly update #4
Mar 9	Spring Break (no class)	
Mar 16	Weekly standup #5 Inference review	Weekly update #5
Mar 23	Weekly standup #6 Report writing	Weekly update #6
Mar 30	Weekly standup #7 Plagiarism, Fair use	Weekly update #7
Apr 6	Weekly standup #8 Data ethics review	Weekly update #8
Apr 13	Weekly standup #9 Guest appearance	Weekly update #9
Apr 20	Weekly standup #10 Guest appearance	Weekly update #10
Apr 27	Mock Presentation I	Weekly update #11
Apr 28	Makeup day	
Apr 29	Mock Presentation II (Designated Monday)	
May 1	Poster Presentation	
May 3	Video presentation due Report due	

GRADING

Your final grade will be determined by:

- ✓ Project Proposal (5%)
- ✓ Elevator Pitch (5%)
- ✓ Weekly Reports (20%)
- ✓ Standups (10%)
- ✓ Presentation (15%)
- ✓ GitHub Repo (15%)
- ✓ Poster (10%)
- ✓ Data Ethics (5%)
- ✓ Final Report (15%)

Letter Grades

A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
≥ 93	90-92	87-89	83-87	80-83	77-80	73-77	70-73	67-70	63-67	60-63	< 60

DELIVERABLES

1. **Project proposal** (5%): A 1- to 2-page proposal outlining the topic and research questions (which may be revised later) and providing at least an outline of the data sources.
2. **Elevator / 3-min pitch**: (5%) With the research topic selected, try to convince a potential client that your topic is worth pursuing and exciting, or how it is for the social good. Each student will have a maximum of 5 minutes to pitch. If the project is a team's work, each member will need to participate, with a max of 5-minute talk time for each person—no need to use slides (you can, but not required). The time limit is strictly enforced.
3. **Weekly progress**: (20%) Every week, submit a 200-word or less summary of the week's work on Blackboard for grading. The GitHub activities are also an essential part of the grade.
4. **Weekly standups**: (10%) *We must* meet every week.
 - For the second half of the class, we break out into small groups of 3-4 students each to share and discuss our work. In small-group discussions, you can discuss what you have done, what you have learned, and what the barriers are. Give each other comments and suggestions.
 - After each small-group standup, each group will report to the class to summarize its projects. The spokesperson will be randomly chosen.
 - Students will assign a grade to each member of their group at the end (on Blackboard). If a student is absent for more than 50% of class time, you must assign a score of 0. If you do not provide a grade for others, you will not receive a grade yourself.
 - Towards the end of the semester, each student will do a mock presentation. Each day, 5-6 students should present.
5. **Poster Session**: (15%) We are scheduling a dedicated poster session on **Dec 8** at **TBD**. Location **TBD**. This poster can be used for a conference submission or the like. The primary goal is to create a poster that compels viewers to focus on the most important elements of the work. Organize them on a relatively small canvas. It surely will help with preparing for the final presentation.

The GWU library has a CREATE digital studio department (go.gwu.edu/create). The GWU community can use the service free of charge. It includes 3D printing (150mm wide), workstations with the full Adobe Creative Suite and other design software, a GoPro camera available for loan, a studio for creating podcasts and other digital content, and, in our case, a large-format printer up to 42 inches wide. All are free of charge (with per-semester limits). You will need to create a poster measuring 36" x 48".

6. **Video presentation:** (15%) Each student/team will deliver a video presentation of their project at the end of the semester. Students should produce a 15-minute presentation for individual projects. For teams with 2 team members, a 25-minute presentation, and for teams with 3 members, a 30-minute presentation. After I review your presentations, I may need to schedule you/your team for a Q&A session within 5 days of the due date.
7. **GitHub:** (15%) All your work should be kept on a (private) GitHub Repo. Follow the given template and put your work there.
8. **Final Report:** (15%) A final written report to be submitted. You can incorporate feedback and suggestions received during the presentations. Please note that GWU has a robust Writing Center to assist students with a wide range of writing tasks. Please consider using it early, rather than in the last few days.

GENERATIVE ARTIFICIAL INTELLIGENCE (GAI) IN OUR COURSE

I agree that the use of GAI is inevitable and is how modern programmers work. Not allowing you to use GAI when coding is like not allowing someone to use a spell checker when writing, or not allowing someone to use an automobile when commuting. I simply ask for two things when you do use it:

1. Cite it as you would for other references. Indicate where you used GAI for your code. Sample citation: "Create a script in Python to web scrape NBA.com season data" prompt. ChatGPT-4, OpenAI, 14 Mar. 2023, chat.openai.com/chat.
2. Thoroughly understand the code you use in your work. Rewrite/restructure it so that it makes sense to you. Seek help from yourself or others if you do not understand the code you are using. I will ask you what your codes mean and why.

Maintaining academic honesty is a challenging skill to learn. There are also some University-wide guidelines you need to be aware of:

- GW Office of the Provost - Generative Artificial Intelligence guidelines
- An aggregated source for GAI classroom policies
- Responding to AI tools when designing activities

That said, you should also be aware of the limitations and ethical issues associated with AI tools. These tools are based on training data that is likely biased, incomplete, and outdated, to name a few obvious deficiencies. If we base our everyday decisions on all the books and articles that were written hundreds of years ago, for example, there would be no women's rights, and no ethnic equality in our decisions. I am confident you will delve into more of these questions in another data science and ethics class in the future.

UNIVERSITY POLICIES

OBSERVANCE OF RELIGIOUS HOLIDAYS

Students must notify faculty during the first week of the semester in which they are enrolled in the course, or as early as possible, but no later than three weeks prior to the absence, of their intention to be absent from class on their day(s) of religious observance. If the holiday falls within the first three weeks of class, the student must inform the faculty

in the first week of the semester. For details and policy, see “Religious Holidays” at provost.gwu.edu/policies-procedures-and-guidelines.

ACADEMIC INTEGRITY CODE

Academic integrity is an essential part of the educational process, and all members of the GW community take these matters very seriously. As the instructor of record for this course, my role is to provide clear expectations and uphold them in all assessments. Violations of academic integrity occur when students fail to cite research sources properly, engage in unauthorized collaboration, falsify data, and otherwise violate the Code of Academic Integrity. If you have any questions about whether or not particular academic practices or resources are permitted, you should ask me for clarification. If you are reported for an academic integrity violation, you should contact the Office of Student Rights and Responsibilities (SRR) to learn more about your rights and options in the process. Consequences can range from failure of the assignment to expulsion from the university and may include a transcript notation. For more information, please refer to the SRR website (<https://studentconduct.gwu.edu/academic-integrity>), email rights@gwu.edu, or call 202-994-6757.

USE OF ELECTRONIC COURSE MATERIALS AND CLASS RECORDINGS

Students are encouraged to use electronic course materials, including recorded class sessions, for private personal use in connection with their academic program of study. Electronic course materials and recorded class sessions should not be shared or used for non-course-related purposes unless express permission has been granted by the instructor. Students who impermissibly share any electronic course materials are subject to discipline under the Student Code of Conduct. Please contact the instructor if you have questions regarding what constitutes permissible or impermissible use of electronic course materials and/or recorded class sessions. Please contact Disability Support Services at disabilitysupport.gwu.edu if you have questions or need assistance in accessing electronic course materials.

ACADEMIC SUPPORT

WRITING CENTER

GW’s Writing Center cultivates confident writers in the University community by facilitating collaborative, critical, and inclusive conversations at all stages of the writing process. Working alongside peer mentors, writers develop strategies to write independently in academic and public settings. Appointments can be booked online at gwu.mywconline.com.

ACADEMIC COMMONS

Academic Commons provides tutoring and other academic support resources to students in many courses. Students can schedule virtual one-on-one appointments or attend virtual drop-in sessions. Students may schedule an appointment, review the tutoring schedule, access other academic support resources, or obtain assistance at academiccommons.gwu.edu.

SUPPORT FOR STUDENTS OUTSIDE THE CLASSROOM

DISABILITY SUPPORT SERVICES (DSS) 202-994-8250

Any student who may need an accommodation based on the potential impact of a disability should contact Disability Support Services at disabilitysupport.gwu.edu to establish eligibility and to coordinate reasonable accommodations

COUNSELING AND PSYCHOLOGICAL SERVICES 202-994-5300

GW's Colonial Health Center offers counseling and psychological services, supporting mental health and personal development by collaborating directly with students to overcome challenges and difficulties that may interfere with academic, emotional, and personal success. healthcenter.gwu.edu/counseling-and-psychological-services.

SAFETY AND SECURITY

- In an emergency: call GWPD at 202-994-6111 or 911
- For situation-specific actions: review the Emergency Response Handbook at: safety.gwu.edu/emergency-response-handbook
- In an active violence situation: Get Out, Hide Out, or Take Out. See go.gwu.edu/shooterprep
- Stay informed: safety.gwu.edu/stay-informed