Syllabus

Tuesday Sep 9, 2025 at 03:08 pm

> **!** Important
>
> **You must familiarize yourself with this document. It is thorough and covers our expectations and policies.**

Instructor

- Abhijit Dasgupta

Communication

> **!** Important
>
> The primary mode of communication will be **Slack**. The course's Slack workspace is https://enia5120fall2025.slack.com.. You will receive an invitation to join.
> **Instructor e-mail:** abhijit.dasgupta@georgetown.edu. This is the preferred way of contacting the professor privately. Please read the communication and Slack rules.

Course description

This course provides an introduction to data science where we will explore the principles, methods, and tools that drive modern data-driven inquiry, inference, prediction and decision-making. This course is designed to provide you with a strong foundation in the core elements of the field, including data wrangling, exploratory analysis, statistical modeling, machine learning, and effective data visualization. Fundamental to this is an conceptual and practical understanding of uncertainty, association and causality. We will emphasize both the theoretical underpinnings and the practical applications of data science, preparing you to critically

evaluate data, build robust analytical workflows, and effectively communicate results to diverse audiences. Through hands-on exercises, real-world datasets, and collaborative projects, you will gain not only technical proficiency but also the problem-solving mindset that defines successful data scientists.

This course will primarily use the statistical analytics program R and the GUI-based analytics platform JASP which runs on R for analyses, simulations, visualizations, and enhancing statistical understanding.

Learning objectives

- Develop a foundational understanding of how to translate data into insights, inference and predictions
- Understand the central role of uncertainty and how it influences how we interpret data
- Think critically about data, including the appropriateness of study design, representativeness, and what you can really answer from a particular data set.
- Appreciate issues like bias and confounding that can affect our inferences from data
- Use high-quality data visualizations to tell your story
- Effectively use statistical software to ingest, transform, analyze, visualize and report on data in the appropriate context
- Communicate findings effectively using compelling visualizations and clear, audience-appropriate narratives.
- Collaborate on data-centric projects, developing teamwork and project management skills essential in professional data science environments.

Pre-requisites

- There are no formal pre-requisites, though an prior introduction to algebra will be very helpful
- Familiarity with R will be helpful

We will encourage the use of *version control systems* like **git** in this class. Some tutorials to brush up on these skills:

- git - the simple guide
- Nico Riedmann's Learn git concepts, not commands

Required resources

Computer

You should have a laptop (no Chromebooks, please). Windows, Mac or Linux machines are acceptable. Please bring your machine to class, since we will be having interactive activities and labs.

Paper and writing/drawing implements

You should have access to blank paper and writing/drawing instruments. A pad/notebook and a pen/pencil will suffice.

Evaluation

- Attendance: 10%
- Engagement and active learning: 20%
- Assignments: 30%
- Lab completions and quizzes: 10%
- Project: 30%

Learning activities

Class format

Each weekly class will comprise approximately 90 minutes of lecture and discussion and approximately 60 minutes of computational or data-driven laboratory. Slides will drive each lecture, but they will serve mainly as a framework and reference material for the class. The expectation is that students will come prepared to engage and ask questions, since much of the semester will be focused on foundational concepts that require thinking and understanding. Questions and discussion are the gateway to understanding and intuition.

Readings and pre-work

Readings, case studies and other pre-work will be assigned every week to introduce you to the topic of the week. It is **essential** that you complete this prior to class and are prepared to participate and engage in class. These materials will be made available on Canvas and on the website.

> **!** Important
>
> You must read the readings and complete the pre-work prior to class time so you are prepared to participate in class.

Case studies and laboratory work

We will discuss concepts, theory and methods most weeks, and have a laboratory section discussing both conceptual and technical aspects every week. These laboratory sessions will use JASP initially, and then R.

> **!** Important
>
> - You will start the labs in class but you will most likely not finish. It is your responsibility to complete the labs to enable your learning. **Completing the labs succesfully is also part of your grade.**

Engagement and participation

We don't expect merely token participation (just asking simple or rhetorical questions) but high quality content in the engagements. You are all high-quality students, and this is the expectation we carry. There are several ways we would like to interact with you:

- Participation during class, asking questions, clarifying doubts, and even answering your classmate's questions
- Each class will have a discussion on Canvas, for each section. Each student will do a "muddy water" exercise after each lecture, highlighting in short paragraphs (a) one thing that clicked for you in class, (b) one thing that remains confusing (especially conceptual topics), and (c) one thing that you think might have been done in class to improve your understanding
- Each lecture will have an associated Google document that everyone can edit. The intent is to have a shared set of notes generated by you so that we can co-learn and co-create through class. These will be living documents where we all share our understanding of the topic(s) at hand.

We will keep track of who is contributing to the class as a baseline evaluation, but full credit can be obtained with higher quality questions, comments, independent ideas, evidence of analytic thinking, and insights that help us all understand. We're really looking forward to this aspect of the class.

> **!** Important
>
> - **Classes and labs will not be recorded.** The best way to learn and benefit from this class is in-person participation.

Homework assignments

There will be several homework assignments. The goal of these problem sets is to solidify concepts and learning objectives each week, as we build skills and knowledge throughout the semester These will typically involve a set of questions to reinforce the conceptual material of the week, and a set of coding/analysis exercises to solidify their implementation. Submissions will typically be done on Canvas. We will introduce Quarto as the primary tool for writing and submitting your work, which will allow us a single means of both textual and coding submissions.

The intent of the homework is to reinforce concepts, but not be overly burdensome. These will generally be thinking exercises. A grading rubric will be provided for all assignments.

Homework assignments will be published after class, and will be due the following Sunday by 11:59pm Eastern time (but see the calendar for any deviations)

> **!** Important
>
> Familiarize yourself with the submission requirements and the late policy for this class.

> **i** Note
>
> Since this is a graduate-level class, we expect students to want to learn and not search online for answers. See the Academic Integrity section for more details.

Online Quizzes

We will often have quizzes to ensure that the pre-work material has been absorbed. *These will be graded for completion.*

There will be unannounced quizzes a few times during the semester, at random intervals and times. The quizzes ensure you are keeping up with the material presented in the class. The material for the quizzes will be drawn from lectures, labs, and readings.

> **!** Important
>
> Missed quizzes cannot be made up.

Project

You will assemble into groups of 2 to 3 students. The objective is to have 10 groups in total. Over the course of the semester you will work towards creating an analytic project relevant to the program of study that will incorporate visualization, inference and prediction through the use of data science techniques. Details will be available in the project website. by week 2.

> **!** Important
>
> The project will have several milestones that are cumulative in nature. Therefore, we will grade the project after the final submission with a holistic project rubric. We will grade the milestones in a qualitative way, and we will provide feedback and a trending grade with each milestone. It is up to you to incorporate the feedback provided. If your milestone trending grade is lower than you expected, and you do not incorporate the feedback we provide for improvement, do not expect your final project grade to improve.

Total is 100 points and the final grade will be based on the scale below. There is no plan to curve the final grade.

- A: >= 95
- A-: >= 90, < 95
- B+: >= 87 < 90
- B: >= 83, < 87
- B-: >= 80, < 83
- C: >= 70, < 80 (last passing grade)
- F: < 70

> **!** Important
>
> Failing this course is highly unlikely but definitely possible. Reasons for failing include but are not limited to:
>
> - Consistently delivering work that is significantly below expectations
> - Consistently missing deliverables
> - Consistently missing class
> - Being found in violation of academic integrity

Grading philosophy

Some of the assignments you will work on are open-ended and some are not (i.e. specific tasks). Grading is generally holistic, meaning that there may not always be specific point value for individual elements of a deliverable. Each deliverable submission is unique and will be compared to all other submissions.

Deliverables that:

- **Exceed the requirements and expectations** are typically considered A level work.
- **Just meet the requirements and expectations** are typically considered B to A- level work.
- **Do not meet the requirements** are typically considered B- or lesser level work.

Partial credit will be given where appropriate.

All deliverables must meet general quality requirements that are expected from students at the graduate school level as well as specific requirements that will be provided for each deliverable.

> **!** Important
>
> Portions of each assignment may be programmatically graded. To facilitate this, it is **essential** that folder structures and file names exactly match what the provided directions state. Failure to do so makes grading more difficult and time-consuming, especially in a class with over 100 students, and will lead to partial point deductions.

Points **will** be deducted for any of the following reasons:

- Your technical approach is fundamentally flawed
- Your analytical decisions are unjustified
- You did not follow any direct and specific instructions
- You do not have proper spelling and grammar in the textual parts of the assignment
- Your deliverable has missing sections
- Your overall presentation and/or writing is sloppy
- You use incorrect file names (wrong extensions, wrong case, etc.)

Submitting your work

All submissions will be via Canvas. The preferred method will be as a Quarto document that has been rendered into PDF or HTML.

Late policy

For homework assignments, if an extension past the deadline is requested, the following rules will apply:

- A late penalty of 10% per day, up to 4 days, will be assessed for assignments and labs that are submitted after the deadline. You may still submit a missed lab or assignment up until the last day of classes with a maximum possible grade of 50%. You are encouraged to do so to receive feedback.
- Missed in-class quizzes cannot be made up and will receive a grade of zero.
- Project deadlines are fixed and have no extensions or late penalty. A missed project deliverable will receive a grade of zero.

Other course policies

Attendance and punctuality

Attendance is mandatory and will be taken. Please be punctual and respectful and remain for the full period until excused.

Given the technical nature of this course, and the breadth of topics discussed, you are expected to attend each class, to complete all readings, and to participate actively in lectures, discussions and exercises. We understand there may be times you may need to miss class; please inform us in advance if you are not able to attend class for any reason so we can provide class materials and also prepare for a smaller classroom experience. However, it is up to you to keep up.

Please don't take advantage of my sympathy and empathy. There are legitimate reasons for missing class, including illness, family emergencies, and unavoidable work situations. Let me know as soon as you can if you will not be able to attend class.

Participation

We love participation. Read. Raise your hand. Ask questions. Make comments. Challenge us. Acknowledge us. If we speak for three hours to a silent classroom, it is a lot more boring and tiring for everyone.

Laptop and phone use

You must bring your laptop to class to work on labs. No phone use is allowed during lecture.

We would prefer that during class time, you use your laptops for class-specific activities, including analyses, on-the-fly searches for understanding and note-taking. This class is not easy, so your attention and focus will help you succeed.

Communication and Slack Rules

- All announcements will be posted on Canvas and Slack
- Use Slack for any question you may have about the course, about assignments or any technical issue. This way everyone can learn from each others questions. We will be monitoring and providing answers on a regular basis. **Make sure you understand what is allowed in Slack.**
- All communication with the professor should be either on general Slack channels or via the abhijit.dasgupta@georgetown.edu email. *. E-mail is preferred for any communication of a personal nature as opposed to questions about class materials.
  - Slack DMs are not to be used **unless** we DM you first and you can respond to our message. Students may not initiate DMs.
  - Any email sent to the instructor containing any course question that is not personal in nature will not be answered; these are expected to go on Slack for general discussion and peer contributions.
  - E-mails of a personal nature (personal issues or difficulties, questions regarding grades, seeking help with a personal issue, etc.) should be sent to the instructors' e-mail (abhijit.dasgupta@georgetown.edu).
- Keep an eye on the questions posted in Slack. Use the search function. It's very possible that we have already answered a question, and we reserve the right to point you to the syllabus, previous Slack messages, or other document containing the information requested.
- Assignment, lab and project questions will only be answered on Slack up to 12 hours before something is due.

Open Door Policy

Please approach or get in touch with me if something is not working for you regarding the class, methods, etc. My pledge to you is to provide the best learning experience possible. If you have any issue please do not wait until the last minute to speak with us. I care deeply about your learning and success.

If you are struggling in any way, reach out so we can consider accommodations to improve your experience. Let me know early and often how you're doing. I will also be checking in. There are many resources that you can avail if you are going to physical or mental issues. Please reach out and I'm ready to help.

Academic Integrity

As a Jesuit, Catholic university, committed to the education of the whole person, Georgetown expects all members of the academic community, students and faculty, to strive for excellence

in scholarship and in character.The University spells out the specific minimum standards for academic integrity in its Honor Code, as well as the procedures to be followed if academic dishonesty is suspected.

Over and above the honor code, in this course we will seek to create an engaged and passionate learning environment, characterized by respect and courtesy in both our discourse and our ways of paying attention to one another.

The code of academic integrity applies to all courses at Georgetown University. Please become familiar with the code. All students are expected to maintain the highest level of academic integrity throughout the course of the semester.Please note that acts of academic dishonesty during the course will be prosecuted and harsh penalties may be sought for such acts. Students are responsible for knowing what acts constitute academic dishonesty. The code may be found at https://bulletin.georgetown.edu/regulations/honor/.

> Caution
>
> **We have a ZERO TOLERANCE POLICY and students found to be in violation will be reported and penalized. The consequences of any violation may include: additional points penalty, getting a grade of zero, automatically failing the course, and suspension or expulsion from the program.**

Definition of collaboration

In the spirit of fostering a collective and inclusive learning environment, we acknowledge that you will work and study with your peers. We also acknowledge that you use web resources (code examples specifically), and that in writing a program many of you will most likely use the same libraries, functions and other similar instructions in your scripts. However:

- **You must write your own code.** This will be verified for every assignment against every submission, and any similarity greater than 60% between students on a given assignment will be considered to be unauthorized collaboration.

What is allowed

- Collaborating with other students during in-class labs to facilitate collective learning
- Using Slack for helping one-another as long as:
  - You do not provide answers directly but only discuss potential approaches
  - You only share up to a few lines of code for everyone's benefit for the resolution of a specific question or issue
- Using anything (code, resources, tips, approaches, etc.) provided by the instructional team

What is forbidden

The following actions are **not permitted in any way** and are considered a violation of academic integrity:

- Copying and sharing code *between students* in **individual** assignments or *across goups* in the **group project**
- Sharing anything on any individual assignment
- Using code snippets found online (stack overflow, etc.) and not commenting the source
- Plagiarism of any kind
- Using any Generative Artificial Intelligence tool without acknowledging it
- Making your private GitHub repos public
- Sharing or posting any course materials anywhere
- Faking or tampering with git commit dates or messages

Use of Generative AI tools

We recognize the recent availability of very powerful generative AI tools like Chat-GPT, GitHub Copilot, and others. These tools can help us be more effective and we embrace their use.

However (and you know there's a 'however'), using GAI tools *in place of* thinking and gaining experience is not fruitful for you and will not serve you well. Intuition and understanding is gained by thinking, ruminating and discussing, not by the short-term objective of completing assignments and getting points. The concepts and tools we are exploring in this class will affect most aspects of your future career, and learning them well will benefit you. Inappropriate use of data and data scientific tools results in a "garbage in, garbage out" situation that can affect decisions and can cost time and money. So learning this stuff instead of relying on a hallucination-prone GAI tool at this foundational stage will probably be a good thing^{TM}.

> **!** Important
>
> You are allowed to use GAI tools in a *non substantial* way.
> What does *non substantial* mean?
> **It means that whatever is generated by GAI must not make up the majority of the work you do.**
> Any use of these tools must abide to the following rules:
>
> - You must acknowledge the use of GAI tools
> - You must comment which code blocks were generated by GAI
> - You must note which written sections were generated by GAI
> - If you used a prompt to ask the GAI tool to do something, you must include it

For this course, valid uses of gen-ai can be:

- Generating a code snippet or single function to perform a task. It's likely you'll need to modify it anyway
- Commenting code
- Using it as a writing aid (spelling, grammar, word choice, limited phrase translation) on content created by you, not the actual writing. **Note:** non-native English speakers cannot use gen-ai to fully translate content written in another language.
- Using it to assimilate openly available resources (papers, blogs, etc) that can allow you a launching pad to understand a topic. However, it is a case of "buyer beware" in terms of the reliability and veracity of the sources that ge-ai uses.

⚠ Warning

Any deviation from these rules is considered a violation of academic integrity and will be acted on.
You typically KNOW when you are crossing the line into un-ethical territory. As a general rule, If you feel like you might be crossing a line, then you probably are!

Georgetown University resources and policies

Georgetown University's Plagiarism Policy

Plagiarism or academic dishonesty in any form will not be tolerated and may result in a failing grade. All Honor Code violations will be submitted to the Honor Council.

Academic integrity is central to the learning and teaching process. Students are expected to conduct themselves in a manner that will contribute to the maintenance of academic integrity by making all reasonable efforts to prevent the occurrence of academic dishonesty. Academic dishonesty includes (but is not limited to) obtaining or giving aid on an examination, having unauthorized prior knowledge of an examination, doing work for another student, and plagiarism of all types, including copying code.

Plagiarism is the intentional or unintentional presentation of another person's idea or product as one's own. Plagiarism includes, but is not limited to the following: copying verbatim all or part of another's written work; using phrases, charts, figures, illustrations, code, or mathematical/scientific solutions without citing the source; paraphrasing ideas, conclusions, or research without citing the source; and using all or part of a literary plot, poem, film, musical score, or other artistic product without attributing the work to its creator. Students can avoid unintentional plagiarism by following carefully accepted scholarly practices. Notes taken for papers and research projects should accurately record sources cited, quoted, paraphrased, or summarized sources and articles should be acknowledged in footnotes.

Honor System

All students are expected to maintain the highest standards of academic and personal integrity in pursuit of their education at Georgetown. Academic dishonesty, including plagiarism, in any form, is a serious offense, and students found in violation are subject to academic penalties that include, but are not limited to, failure of the course, termination from the program, and revocation of degrees already conferred. All students are held to the Georgetown University Honor Code. For more information about the Honor Code http://gervaseprograms.georgetown.edu/honor/

Academic Integrity and Courtesy

As a Jesuit, Catholic university committed to the education of the whole person, Georgetown expects all members of the academic community, students and faculty, to strive for excellence in scholarship and in character. The University spells out the specific minimum standards for academic integrity in its Honor Code and the procedures to be followed if academic dishonesty is suspected. Over and above the honor code, in this course, we will seek to create an engaged and passionate learning environment characterized by respect and courtesy in both our discourse and our ways of paying attention to one another.

Academic Resource Center

The Academic Resource Center (ARC) is the campus office responsible for reviewing medical documentation and determining reasonable accommodations for students with disabilities. You can reach the ARC via email at arc@georgetown.edu.

Counseling and Psychiatric Services (CAPS)

As Georgetown faculty, you are among the most important individuals in some of the students' lives. They may turn to you when they are struggling and in times of need, or you may be one of the first to notice when they are distressed.

The CAPS website has tips for faculty on how to deal with struggling or distressed students. 202.687.6985 or after hours, call (833) 960-3006 to reach Fonemed, a telehealth service; individuals may ask for the on-call CAPS clinician.

Emergency Preparedness and HOYAlert

We encourage all faculty to become familiar with Georgetown's Office of Emergency Management and sign up for HOYAlert to receive important safety and University operating status updates. Faculty teaching at the Georgetown Downtown campus might also want to sign up for AlertDC to obtain safety and traffic updates.

Office of Institutional Compliance and Ethics

The Office of Institutional Compliance and Ethics supports and coordinates many compliance-related activities the University undertakes. With the endorsement and assistance of the University's senior leadership, this Office is responsible for leading the development, implementation, and operation of the Georgetown Institutional Compliance and Ethics Program.

Office of Institutional Diversity, Equity and Affirmative Action (IDEAA)

The mission of IDEAA is to promote a deep understanding and appreciation among the diverse members of the University community to result in justice and equality in educational, employment, and contracting opportunities, as well as to lead efforts to create an inclusive academic and work environment.

Title IX/Sexual Misconduct

Georgetown University and its faculty are committed to supporting survivors and those impacted by sexual misconduct, which includes sexual assault, sexual harassment, relationship violence, and stalking. Georgetown requires faculty members unless otherwise designated as confidential, to report all disclosures of sexual misconduct to the University Title IX Coordinator or a Deputy Title IX Coordinator. Suppose you disclose an incident of sexual misconduct to a professor in or outside of the classroom (except disclosures in papers). In that case, that faculty member must report the incident to the Title IX Coordinator or Deputy Title IX Coordinator. The coordinator will, in turn, reach out to the student to provide support, resources, and the option to meet—[Please note that the student is not required to meet with the Title IX coordinator.]. More information about reporting options and resources can be found on the Sexual Misconduct Website.

If you would prefer to speak to someone confidentially, Georgetown has a number of fully confidential professional resources that can provide support and assistance. These resources include:

- Health Education Services for Sexual Assault Response and Prevention: confidential email sarp@georgetown.edu

- Counseling and Psychiatric Services (CAPS): 202.687.6985 or after hours, call (833) 960-3006 to reach Fonemed, a telehealth service; individuals may ask for the on-call CAPS clinician

Title IX Sexual Misconduct Statement Please know that as faculty members, we are committed to supporting survivors of sexual misconduct, including relationship violence and sexual assault. However, university policy also requires us to report any disclosures about sexual misconduct to the Title IX Coordinator, whose role is to coordinate the University's response to sexual misconduct.

Georgetown has a number of fully confidential professional resources who can provide support and assistance to survivors of sexual assault and other forms of sexual misconduct. These resources include:

- Getting Help
- Jen Schweer, MA, LPCAssociate Director of Health Education Services for Sexual Assault Response and Prevention (202) 687-032jls242@georgetown.edu
- Erica Shirley, Trauma SpecialistCounseling and Psychiatric Services (CAPS)(202) 687-6985els54@georgetown.edu

Threat Assessment

Georgetown University established its Threat Assessment program as part of an extensive emergency planning initiative. The program at Georgetown has been developed and implemented to meet current best practices and national standards for hazard planning in higher education institutions and workplace violence prevention.

Special Accommodations

If you believe that you have a disability that will affect your performance in this class, don't hesitate to get in touch with the Academic Resource Center for further information. The center is located in the Leavey Center, Suite 335. The Academic Resource Center is the campus office responsible for reviewing documentation provided by students with disabilities and determining reasonable accommodations according to the Americans with Disabilities Act (ADA) and University policies.