



DATA SCIENCE & AI FOR BUSINESS
(TECHNICAL)
TECH-GB.2336
Wednesdays 6pm-9pm

SYLLABUS – Spring 2026

Instructor	Foster Provost , Technology, Operations and Statistics Department
Office; Hours	FP: KMC 8-86; Office hours: Tues 5pm-6pm and by appt.
Email	fprovost@stern.nyu.edu Begin subject: [DS Class] ... ← note!
Classroom	TBD
Class time	Wed 6pm-9pm
First : Last Class	Feb 4; April 29
Teaching Fellow	Baotong Zhang; bz2274@stern.nyu.edu; Office: KMC TBD Hours: TBD

NOTE: It is required that you have programmed in Python prior to taking this class. That does not mean that you have to be a Python expert or an expert programmer. An introduction to Python class or serious bootcamp should suffice (take one over the holiday?). We will not be teaching either Python or programming in this class, and you likely will not have time to learn from scratch on the fly. You will learn the data analysis and machine learning libraries that we will use (mainly Pandas and sklearn). You will be able to use GenAI to help you with the coding (see policy below).

1. Course Overview

This course is about the data science way of thinking about data and its role in business, and about the fundamentals of machine learning and building AI-driven solutions.

We will examine how advanced analytics and machine learning can be used to improve decision-making. We will study the fundamental principles and techniques of data science, machine learning, and data-driven artificial intelligence, and we will examine real-world examples and cases to place data science/AI in context, to develop data-analytic thinking, and to illustrate that proper application is as much art as science.

After taking this course you should:

1. *Approach business problems data-analytically.* Think carefully & systematically about whether & how data and ML/AI can improve business outcomes, to make better-informed decisions for product management, marketing, sales, operations, consulting, etc.
2. *Be able to interact competently about data science, machine learning, and data-driven artificial intelligence.* Know the fundamental principles of data science that underlie processes, algorithms & systems. Improve your ability to envision new opportunities, work on data science projects & interact with everyone involved.
3. *Have had some hands-on experience building and evaluating machine learning models.* This is very helpful if you will be interacting closely with data/ML scientists/engineers, especially on AI/ML products and services. The respect you'll garner will be invaluable for your relationship. Some of you will be prepared to pilot ideas or proof-of-concept opportunities that present themselves.

2. Focus and interaction

The course will explain the fundamental principles, uses, and some technical details of data science, machine learning, and data-driven AI. We do this through lectures, discussions, in-class case studies, hands-on work in Python, and real-world examples. The emphasis primarily is on understanding and applying the fundamental concepts of data science and business applications of machine learning-based AI. We will discuss the mechanics of how the methods work to a limited degree—only as is necessary to understand and illustrate the fundamental concepts and business applications. This is not an algorithms course. However, many machine learning algorithms and techniques are the embodiment of one or more of the fundamental principles of data science.

We will expect you to be prepared for class discussions by having satisfied yourself that you understand what we have done in the prior classes and in the readings. The assigned readings will cover the fundamental material. The class meetings will be a combination of lectures/discussions on the fundamental material, discussions of business applications of the ideas and techniques, case discussions, student exercises, and demos.

You are expected to attend every class session, to arrive at or before the starting time, to remain for the entire class, and to follow basic classroom etiquette. In general, we will follow Stern default policies unless we state otherwise. We will assume that you have read them and agree to abide by them:

<http://www.stern.nyu.edu/portal-partners/academic-affairs-advising/policies-procedures/default-policies-stern-courses>

The Brightspace site for this course will contain reading materials, assignments, and late-breaking news. You should check the site daily, and we will assume that you keep up to date on all announcements and class discussion.

If you have questions about class material that you do not want to ask in class, or that would take us well off topic, please ask after class, come to office hours to see one of us (Foster, Baotong), or ask in the discussion forum. **All questions related to homework, class material, and technical details must be posted in the discussion forum.** This is to track better your questions and allow other students to benefit from the answers. Also, please try to answer your classmates' questions. In grading your class contribution, we will include your contributions to the discussion forum. You will not be penalized for being wrong in trying to contribute on the discussion forum (or in class).

Worth repetition: It is your responsibility to check Brightspace and your email at least once a day during the week (M-F), and we will assume that you are aware of any announcements that are sent with that timeframe in mind.

Use email for personal issues and for questions specific to your term project. **We will check for class-related email at least once a day. *Your email should include the special tag [DS Class] in the email subject header.*** We use this tag to make sure to process class-related email first. If you do not include the special tag, we may not read the email for a while (maybe a long while). If you forget and send without the tag and then remember later, ***just send it again*** including the tag.

3. Readings

Book: The textbook for the class will be the following.

Data Science for Business: What you need to know about data mining and data analytic thinking. Provost & Fawcett **2020 update**.

This updated version will be available in the bookstore as a digital coursepack. I'll post the details to Brightspace when they send them to me.

This book covers the fundamental material that will provide the basis for you to think and communicate about data science and business analytics. We will complement the book with class notebooks, plus discussions of applications, cases, and demonstrations, and possibly additional readings for material that is not covered in the book. If any of those readings are required, we will note that clearly. "Supplemental" readings are optional.

You are expected to ask questions about any required material that is unclear after our class discussion and after reading the assigned book chapters. Requiring readings from the book frees up class time for more discussion of applications, cases, etc.—so many of your questions may be answered in the book. If not, please let us know! If the book discussion is not adequate to explain a topic we don't cover in class, you should ask about it in the discussion forum. We will be happy to follow up.

Note that some of readings may be accessible for free only from an NYU computer. If you can't access a link from home, please try it from school or ask a classmate.

For those interested in going deeper on the "hands-on" component of the class:
(OPTIONAL)

Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd Edition
by Sebastian Raschka & Vahid Mirjalili
<https://www.amazon.com/Python-Machine-Learning-scikit-learn-TensorFlow/dp/1789955750/>

4. Requirements and Grading

The grade breakdown is as follows:

1. Homeworks: 30%
2. Term Project: 30%
3. Class Contribution/Professionalism/Attendance: 10%
4. Final Quiz: 30%

At NYU Stern we seek to teach challenging courses that allow students to demonstrate differential mastery of the subject matter. Assigning grades that reward excellence and reflect differences in performance is important to ensuring the integrity of our curriculum. In my experience, students generally become engaged with this course and do excellent or very good work, receiving As and Bs, and only a few perform only adequately or below and receive C's or lower. Note that the actual distribution for this course and your own grade will depend upon how well you actually perform this particular semester.

Homework Assignments

The homework assignments are listed (by due date) in the class schedule below. Each homework comprises questions to be answered and/or hands-on tasks. Except as explicitly noted otherwise (see next paragraph), you are expected to complete your assignments on your own—without interacting with others on the completion of your assignment. You are free of course to discuss the concepts with your classmates, and to discuss similar problems to the ones in the homeworks.

For the hands-on parts of the assignments, if they are challenging for you I encourage you to work with your group members and other classmates to understand how to get Python to do what you need to do generally (not on the specific homework problems), and then to complete your assignment on your own. So, for example, you could have a classmate help you do something similar, such that then you would be able to complete the assignment.

With the support of me, Baotong, and your classmates, we operate under a “**diligent attempt but limited frustration**” policy: (1) If you get stuck on something, spend some time to try to find the answer (you can Google it or even use an AI coding assistant—see below). If you seem to be moving forward, keep going. That search and discovery will pay off, both in terms of the direct learning about how to do what you need to do, and also in terms of your learning *how to figure such things out*. BUT, (2) limit frustration—start your assignments early enough that if you run into a wall, you can just stop searching and ask about it. Let’s say, if you feel like you have not moved forward after 15 minutes of being stuck, just stop and ask your classmates or in the discussion forum. If you don’t get a resolution, escalate it to Baotong or me.

Policy on Generative AI: Generative AI (GenAI) tools such as ChatGPT, Gemini, and Copilot have become an important skill in business. You can use of these and other AI tools to aid in coding, idea generation, and overcoming other obstacles in your work. Specifically for coding, these tools can greatly help the speed of getting things done. Please use these powerful tools wisely! If you really want to learn the material, the best approach is to attempt problems yourself, and reach for the power tools when you hit a stumbling block. (Research suggests that AI tools can inhibit learning.) In written examples like homeworks and your term project, you can use GenAI to make suggestions for your coding and writing, but not to do the assignment for you or to write your writeups. Assignments that are deemed to be largely cut-and-paste from GenAI may be marked down. Note that you will not have access to GenAI for the final quiz, so you’ll want to make sure you learn the material (if you want to do well on the final).

Completed assignments must be handed in as specified in class, by the due dates and times we provide. Assignments will be graded and returned promptly. Answers to homework questions should be well thought out and communicated precisely and professionally, avoiding sloppy language, poor diagrams, and irrelevant discussion.

The hands-on tasks in the homeworks will be based on data that we will provide. You will mine the data to get hands-on experience in formulating problems and using the various techniques discussed in class.

IMPORTANT: You must have access to a computer on which you can access the class Colab notebooks. If you do not have such a computer, please see me

immediately so we can make alternative arrangements. If you need additional help with Colab or Python, please see Baotong.

Generally Baotong (our teaching fellow; see first page) should be the first point of contact for questions about and issues with the **homeworks, hands-on work, and Python/Pandas/sklearn**. Baotong will have the responsibility to make sure that all questions are answered in a timely fashion.

Foster will be the main contact for the course material, the class sessions, and the term project deliverables (but Baotong may be able to provide better help than Foster for advice on coding-related questions).

Late Assignments

Assignments are to be submitted by the due date. Assignments up to 24 hours late will have their grade reduced by 15%; assignments up to one week late will have their grade reduced by 40%. After one week, late assignments will receive no credit. Please turn in assignments early if there is any uncertainty about your ability to turn it in on time.

Term Project

Term project deliverables, including the final report, will be prepared by student teams. We will give you the instructions on how to form your teams. Teams are encouraged to interact with Foster and Baotong electronically or face-to-face in developing their projects. You will submit various deliverables through the course. We will discuss the project requirements in class.

Final Quiz

The subject matter covered will be discussed in class.

Participation/Professionalism/Contribution/Attendance

Please see Section 2.

Regrading

If you feel that a calculation, factual, or judgment error has been made in the grading of an assignment or exam, please write a formal memo to me describing the error, within one week after the date on which that assignment was returned. Include documentation (e.g., pages in the book, an entry in one of our notebooks, etc.). I will make a decision and get back to you as soon as I can. Please remember that grading any assignment requires the grader to make many judgments as to how well you have answered the question. Inevitably, some of these go “in your favor” and possibly some go against. *In fairness to all students, the entire assignment or exam will be regraded.*

FOR STUDENTS WITH DISABILITIES: If you have a qualified disability and will require academic accommodation during this course, please contact the Moses Center for Students with Disabilities (CSD, 998-4980) and provide me with a letter from them verifying your registration and outlining the accommodations they recommend. (Usually they contact me, but no harm in checking.) If you will need to take an exam at the CSD, you must submit a completed Exam Accommodations Form to them at least one week prior to the scheduled exam time to be guaranteed accommodation.

Please read the policies for Stern courses

<http://www.stern.nyu.edu/portal-partners/academic-affairs-advising/policies-procedures/default-policies-stern-courses>

Academic Integrity

We take pride in our well-rounded education and approach our academics with honesty and integrity. Indeed, integrity is critical to all that we do here at NYU Stern. As members of our community, all students agree to abide by the [NYU Academic Integrity Policies](#) as well as the NYU Stern Student Code of Conduct, which includes a commitment to:

- Exercise integrity in all aspects of one's academic work including, but not limited to, the preparation and completion of exams, papers and all other course requirements by not engaging in any method or means that provides an unfair advantage.
- Clearly acknowledge the work and efforts of others when submitting written work as one's own. Ideas, data, direct quotations (which should be designated with quotation marks), paraphrasing, creative expression, or any other incorporation of the work of others should be fully referenced.
- Refrain from behaving in ways that knowingly support, assist, or in any way attempt to enable another person to engage in any violation of the Code of Conduct. Our support also includes reporting any observed violations of this Code of Conduct or other School and University policies that are deemed to adversely affect the NYU Stern community.

STERN CODE OF CONDUCT

The Stern Code of Conduct and Judiciary Process applies to all students enrolled in Stern courses.

<https://www.stern.nyu.edu/uc/codeofconduct>.

To help ensure the integrity of our learning community, assignments you submit to NYU Brightspace will be submitted to Turnitin. Turnitin will compare your submission to a database of prior submissions to Turnitin, current and archived Web pages, periodicals, journals, and publications. Additionally, your document will become part of the Turnitin database.

Class Schedule (draft - subject to revision)

Week	Date	Topics <i>(subject to change as class progresses)</i>	Readings	Deliverables
1	Feb 4	<p>Introduction to the Course Introduction to Advanced Analytics Doing Data Science Tools for Advanced Analytics – Conceptual and Practical</p> <p>Notebooks: Dealing with Data w/ Colab, Python & the Command Line</p>	Ch. 1-2	Pre-class Assignment due by 2:30pm TBD
2	Feb 11	<p>Intro to Predictive Modeling Supervised Segmentation Conceptual Tools for Creating a Predictive Analytics Solution to a Business Challenge/Opportunity</p> <p>Notebook: Predictive Modeling/Supervised Segmentation</p>	Ch. 3-4	HW#1 due by 11:59pm TBD
3	Feb 18	<p>Intro to Predictive Modeling, cont. Problem formulation, cont. + Fitting the data. More Modeling</p> <p>Notebook: Fitting mathematical models - Take 1</p>	Ch. 5	Team choices and initial project ideas due by 11:59pm TBD
4	Feb 25	<p>Model performance analytics & the science of predictive modeling Overfitting the data. Holdout testing, cross-validation, learning curves, domain knowledge validation</p> <p>Notebook: Fitting mathematical models - Take 2</p>	Ch. 7	HW#2 due by 11:59pm TBD
5	Mar 4	<p>Model performance analytics II <u>Ranking</u>, true positives, false positives, profit, lift ROC analysis, expected value framework</p> <p>Notebook: Model performance analytics</p>	Ch. 8 Optional: Stein paper	Project Proposal due by 11:59pm TBD
6	Mar 11	<p>Mining fine-grained data, Prediction via evidence combination, possibly: text classification</p> <p>Case Study: Modeling consumer behavior for targeted marketing (banking or online advertising)</p>	Ch. 9 & 10	HW#3 due by 11:59pm TBD
	Mar 18	Spring Break – No Class	None!	

Module	Date	Topics	Readings	Deliverables
7	Mar 25	A/B testing, quantifying uncertainty, confidence intervals, prediction intervals, statistical vs. substantive significance	To Be Posted on Brightspace	
8	Mar 28	From Distributed Representations to Neural Networks + Toward Analytical Engineering	TBP on Brightspace + Ch. 11	HW#4 due by TBD
9	Apr 1	Deep Learning/Large Language Models (like GPT) Modern NLP	TBP on Brightspace	Project update due by 11:59pm TBD
10	Apr 8	ML Strategy I – Which ML projects to invest in?	TBP on Brightspace	
11	Apr 15	Data Science Strategy II, Management & Tactics	Ch. 12 & 13	
12	Apr 22	TBD		
13	Apr 29	<u>Final Quiz</u>		Project report Due by 11:59pm TBD