

Meeting time and location	Day/Time	Location
Lecture 1	Monday 2:10 PM – 3:25 PM	C2-004
Lecture 2	Wednesday 2:10 PM – 3:25 PM	C2-004
Recitation	Friday 3:45 PM – 5 PM	C2-E049
Instructional time requirements are met by the scheduled meeting times.		

Faculty Details	Professor	Instructor
Name	Romain Ferrali	Paul Max Love III
Email	rferrali@nyu.edu	pml9743@nyu.edu
Workspace	A5-L01-1137	A4-157
Office Hours	Tuesday, 5:00 PM – 6:00 PM (Zoom link available on Brightspace)	Thursday, 10:00 AM – 12:00 PM (in person) Friday, 10:00 AM – 12:00 PM (online) Book an appointment

Course Description

Social scientists and policy analysts rely heavily on research drawing on observational data. Students learn to manage and analyze such data and to employ statistical techniques that are common in these applications, with an emphasis on how to translate social science theory into empirical research. Topics include review of basic regression analysis, building multivariate analytical models, and regression analysis with limited dependent variables. The course emphasizes practical training in these skills as well as evaluation, replication, and critical analysis of research conducted in the social science literature. The course is designed as an alternative for students in the social sciences who are not taking Introduction to Econometrics.

Credits: 4

Prerequisites / Corequisite: SOCSC-UH 1010Q or MATH-UH 2011Q

This course counts toward the following NYUAD degree requirements:

- Majors > Political Science > Methods Electives
- Majors > Social Research and Public Policy > Methods Electives
- Majors > Business, Organizations and Society > Methods and Analytics Electives

Course Learning Outcomes and Link to [Program Learning Outcomes \(PLOs\)](#)

Upon successful completion of this course the student will be able to	Linked to Major PLOs
1. Ingest and explore data	2
2. Perform common and important statistical analyses	2
3. Design and produce data visualizations	2
4. Distinguish descriptive analyses from causal analysis	2, 3
5. Assess empirical questions in political methodology and identify the appropriate tools to address them	2, 3

Teaching and Learning Methodologies

- **Lectures:** there will be two lectures per week. Each lecture will begin with 15 minutes of collaborative problem-solving meant to illustrate a key programming technique. Laptops are required in-lecture: we will work together through the lecture to produce code to solve problems related to business analytics, and at the end of each lecture, you will be left with a completed, fully-functional R script to draw from. The last weeks will be devoted to specialized topics, including time-series analysis, geographical and spatial analysis, social network analysis, natural language processing, or advanced data visualization.
- **Recitations:** Recitations are meant to complement and expand course material by focusing on core programming concepts like writing functions and by introducing related tools like LaTeX, Quarto, and R Shiny.
- **Class Participation:** Participation is graded through both quantity and quality. You **do not** get participation credit for attending class. You get participation credit for

participating in class. This includes answering questions in both lectures and recitations, being visibly engaged, and attending office hours.

- **Problem Sets:** There will be 4 problem sets building on material covered during class or in readings. Problem sets will be a mix of programming challenges and applied business analytics, and they are meant to be LOW PRESSURE. They will usually be due one week from the day it is assigned. You are allowed to work in groups, but you are required to write-up your own results and explanations individually and cite anyone you collaborate with. Bottom line: what you submit for your homework should be your own work. If you do work in a group, you are required to list the members of your group at the beginning of your problem set. They will be graded on a binary scale (0/1). In the event that you receive a zero, there is a one-week rewrite policy, assuming that you handed the assignment in on time and made a good-faith effort to complete the assignment. That is, you will have one week to revise and resubmit the assignment in the event you receive a zero. If your submission is still below the bar after resubmission, it will be graded as-is. Most students will have to resubmit most assignments -- this is how we learn! I require that you use LaTeX or Quarto to prepare your homework. This will allow you to make reproducible documents that include both your code and your write-up in one convenient place. Moreover, these systems make documents that look far superior to Word, and manage your bibliography and formatting for you.
- **Midterm practicum:** In a problem set, I provide you with the data, the questions, and what analyses you need to do to answer the questions. In the practicum, I provide the data and the questions, but you'll need to figure out which analyses you need yourself. You will then write a polished report (2000 words max) answering the questions at hand in simple, non-technical terms. The practicum challenges you to apply the material covered in class to a real-world data set based on an open-ended business case study, and write up your analysis and findings. That is, the practicum is an opportunity to actually perform data analytics. We will devote one recitation session to understanding what makes a good practicum response. The case study and data set will be sent to you via email and responses should be submitted via email within two weeks. Students are not allowed to work together or discuss practicum case study solutions under any circumstances. If you have a question about the practicum scenario, send an email. I (or the instructor) will then anonymize your question and respond to the entire class. The goal of the practicum is not to display technical expertise alone, but rather to demonstrate that you can use data and statistical analysis to make an argument regarding a set of questions.
- **Final project:** As the culmination of the course, your final project should be a complete, independent data analysis project. I expect you to find or construct some data, develop a hypothesis, analyze the data to test that hypothesis, produce publication-quality data visualizations, and carefully interpret your results. The project should employ at least one of the methodological tools covered in class to answer a question of substantive interest relating to business analytics using data

assembled by the student. Students should submit a one-paragraph description of the final project and data to be used to the instructor by the end of Spring Break. The proposal should describe the substantive question of interest, a (very) preliminary research design, and discuss the sources of data.

Graded Activities

Activity Detail	Grade Percentage	Submission Date	Linked to Course Learning Outcome(s)
Homework 1	6.66%	2/11	1, 2, 3
Homework 2	6.66%	2/25	1, 2, 3
Homework 3	6.66%	4/20	1, 2, 3, 4, 5
Practicum	25%	4/6	1, 2, 3, 4, 5
Final Project	40%	5/6	1, 2, 3, 4, 5
Participation	15%		

Other Required Readings

- The Effect: An Introduction to Research Design and Causality, by Nick Huntington-Klein

The above book is a free, online textbook. It is available at:

<https://nickchk.com/causalitybook.html>.

Course Schedule

Session	Topic	Reading	Other
1/21	Introductions; syllabus overview; Introduction to R	None	
1/26	Logicals, subsetting, and indexing; the first data set		
1/28	Basic data visualizations; statistics: means and standard deviations	The Effect Chapter 4	

Session	Topic	Reading	Other
2/2	Writing functions; Statistical distributions; Hypothesis testing; Lady Tasting Tea		
2/4	Frequentist hypothesis testing; estimating distributions in R		HW1 Assigned
2/9	Loops; Base R plotting, linear regression	The Effect, 185-201	
2/11	The Tidyverse		HW1 Due
2/16	Interpreting regressions; interaction effects in linear regressions		
2/18	Advanced plots: ggplot2		HW2 Assigned
2/23	String manipulations and regular expressions		
2/25	Polynomial regressions; LOESS regressions	The Effect, 229-236	HW2 Due
3/2	Canceled		
3/4	Merging and reshaping data		
3/9	Canceled		
	Spring break		HW3 Due
3/23	Causality	The Effect, 103-115	Final Project Idea Due

Session	Topic	Reading	Other
3/25	Causality II: causal graphs	The Effect, 117-178	Practicum Assigned
3/30	Cheating at regression	The Effect, 191-196	
4/1	Data Visualization Theory		
4/6	Missing Data		Practicum Due
4/8	Web scraping I		
4/11	Web scraping II		
4/13	Time Series I		HW3 Assigned
4/15	Time Series II		
4/20	Text Data		HW3 Due
4/22	Spatial Data I		
4/25	Spatial Data II		
4/27	Network Data		
4/29	Final Project Workshopping I		
5/4	Final Project Workshopping II		
5/6	Final Project Workshopping III		Final project due May 11

Recitations Schedule. 1/23, 1/30, 2/6, 2/13, 2/20, 2/27, 3/6, 3/27, 4/3, 4/10, 4/17, 4/24, 5/1, 5/8.

Academic Policies: This is a difficult course, and many students find it frustrating. I encourage you to reach out to us if you are struggling with anything. If you spend more than 30 minutes struggling with any single problem, please get email us and we will do our best to gently nudge you in the right direction.

Attendance and Participation: Your participation is expected, required, and graded. This means your physical presence and your full attention. Absence will be penalized. Sleeping in class will be penalized. Lack of attention will be penalized.

Artificial Intelligence: Generative language models like ChatGPT are pretty cool and can be very useful for learning and research. ChatGPT is especially good at many of the basic tasks involved in programming, statistics, and data analysis. Unfortunately, they are not very good for more complicated tasks, and will often mislead you. My goal is for you to learn the course material, and part of that involves struggling with the material. If you rely on these tools too much early on in the course, you will not learn the material and by the end of the course, you will find that ChatGPT cannot help you anymore. This will be bad and you will fail the course. Learning how to use ChatGPT effectively and responsibly for programming, statistics, and data analysis will be integral to this course. You may use it for programming in R during this course but will be held responsible for any mistakes that ChatGPT may have introduced. You may also use it to help you with writing or proofreading.

Integrity: At NYU Abu Dhabi, a commitment to excellence, fairness, honesty, and respect within and outside the classroom is essential to maintaining the integrity of our community. By accepting membership in this community, students, faculty, and staff take responsibility for demonstrating these values in their own conduct and for recognizing and supporting these values in others. In turn, these values create a campus climate that encourages the free exchange of ideas, promotes scholarly excellence through active and creative thought, and allows community members to achieve and be recognized for achieving their highest potential.

Students should be aware that engaging in behaviors that violate the standards of academic integrity will be subject to review and may face the imposition of penalties in accordance with the procedures set out in the NYUAD policy: <https://students.nyuad.nyu.edu/campus-life/student-policies/community-standards-policies/academic-integrity/>

NYU Moses Center for Student Accessibility: New York University is committed to providing equal educational opportunity and participation for students with disabilities. The center works with NYU students to determine appropriate and reasonable accommodations that support equal access to a world-class education. Confidentiality is of the utmost importance. Disability-related information is never disclosed without student permission.

Find further information at:

<https://www.nyu.edu/students/communities-and-groups/students-withdisabilities.html>

Contact: mosescsd@nyu.edu

Mental Health Resources: As a university student, you may experience a range of issues that can interfere with your ability to perform academically or impact your daily functioning, such as: heightened stress; anxiety; difficulty concentrating; sleep disturbance; strained relationships; grief and loss; personal struggles. If you have any well-being or mental health concerns, please visit the Counseling Center on the ground floor of the campus center from 9am-5pm Sunday - Thursday, or schedule an appointment to meet with a counselor by calling: 02-628-8100, or emailing: nyuad.healthcenter@nyu.edu. If you require mental health support outside of these hours call NYU's Wellness Exchange hotline at 02-628-5555, which is available 24 hours a day, 7 days a week. You can also utilize the Wellness Exchange mobile chat feature, details of which you can find on the student portal.