

Regression in Psychology PSYCH-GA.2229

New York University, Spring 2024

Instructor: Prof. Madalina Vlasceanu, Ph.D. [mov209@nyu.edu]
Office: Meyer 505 [office hours by appointment]

Teaching assistant: Kareena del Rosario [kareena.delrosario@nyu.edu]
Office: Meyer 521 [office hours by appointment]

Lecture: Mondays and Wednesdays 2–3pm in Meyer 465 or on [zoom](#)
Lab: Mondays and Wednesdays 3–3:50pm in Meyer 465

Course goals: Students completing this course will have a detailed understanding of multiple regression (MR) as a data-analytic method. Students will review theory and practice of the General Linear Model and learn how MR can be used to carry out analyses of quantitative and categorical data. The relation of MR to correlation, t-tests, ANOVAs, and Mixed Models will be made explicit. Students will solve practical problems in estimating and testing regression models and they will gain experience in carrying out MR analyses using R and Python.

Course Contents: <https://mvlasceanu.github.io/Regression>
https://github.com/kareena-delrosario/regression_rcode

Class Schedule:

Day	Date	Lecture	Lab	Readings	Assignments
1	Jan 22	Lecture 1: Why learn Statistics ?	Installing Python	Ch 1	
2	Jan 24	Lecture 2: Intro to Python	Intro to R	Ch 2	
3	Jan 29	Lecture 3: Data Visualization	Data manipulation in R	Ch 5	
4	Jan 31	Lecture 4: Probability	Data visualization in R	Ch 9	
5	Feb 5	Lecture 5: Correlations in Python		Ch 10	
6	Feb 7	SPSP – NO CLASS			
7	Feb 12	Lecture 6: Hypothesis Testing		Ch 11	
8	Feb 14	Lecture 7: Chi-square test & T-Test	Chi-squares & T-test	Ch 12	
9	Feb 19	PRESIDENTS' DAY – NO CLASS		Ch 13	
10	Feb 21	Lecture 8: ANOVA in Python	ANOVA in R	Ch 14	A.1 due
11	Feb 26	Lecture 9: Model Comparison	ANOVA in R - solved	Ch 15	
12	Feb 28	Lecture 10: Regression in Python	Regression in R		
13	Mar 4	Lecture 11: Regression Assumptions			
14	Mar 6	Lecture 12: Practice Regression	Robust. mod and med		
15	Mar 11	Lecture 13: Mediation in Python		Blair, 2020	A.2 due
16	Mar 13	Lecture 14: Nonlinear Regression		Gureckis, 2021	
17	Mar 18	SPRING BREAK – NO CLASS			
18	Mar 20	SPRING BREAK – NO CLASS			
19	Mar 25	Lecture 15: MixedModelsR / LMMPython		Brown, 2021	
20	Mar 27	Midterm: Practice content covered so far		Bates, 2005	
21	Apr 1	Bayesian Inference (Dr. Joe Bak-Coleman)			A.3 due
22	Apr 3	Power Analysis Simulations (Dr. Jan Voelkel)			
23	Apr 8	Computation (Dr. Felicia Loecherbach)			
24	Apr 10	Network Analysis (Dr. Seungwoong Ha)			
25	Apr 15	Web Scraping (Dr. Ben Guinaudeau)			
26	Apr 17	Stat. Models of Behavior (Dr. Ravi Shroff)			
27	Apr 22	Practical ML (Dr. Michael Morais)			
28	Apr 24	Class project presentation			
29	Apr 29	Class project presentation			

30	May 1	Class project presentation			
31	May 6	Class project presentation			Final Report

Learning Objectives:

- Students will be able to take a dataset and a codebook and conduct valid statistical analyses on different types of variables / data structures to reasonably test different research questions.
- Students will be able to articulate how and why different types of analyses are or are not valid for answering specific research questions with a given set of data.
- Students will be able to critically evaluate scientific articles as to whether the analysis strategy and research design allow for the inferences presented in the paper.

Prerequisites: We assume that all students have successfully completed an intermediate level course in statistical methods covering the basics of statistical inference (including confidence intervals and statistical power), as well as standard univariate tests and techniques.

Requirements and Grading: The final grade will be based on a weighted average of the weekly attendance and class participation (10%), 3 assignments (20%), midterm exam (30%), final report (20%), and final presentation (20% of final grade).

Midterm. One midterm (open-notes) occurs on March 27th that covers all material from lecture and readings up until that point.

Final report: To promote a sense of mastery of regression in the context of actual research, we ask each student to acquire a data set of interest, and to plan, carry out, and report on a complete analysis involving regression. The dataset can be data that you have already collected for your own research, or can be data that is publicly available (e.g., a list of open access datasets can be found [here](#)). However, if using a publicly available dataset, the analysis for this project should be more than a replication of what is already reported. Data sets should be relevant to questions in students' fields and should have more than four key variables. The analysis should consider model specification, examination of the appropriate scale of the variables, outlier and influential data points, handling of missing data, appropriate use of statistical estimation and inference, and power analysis. The final paper should contain two pages that set up the scientific question (i.e., an introduction contextualizing the work), and approximately four pages describing methods and results in a style that would be appropriate for a journal submission. The paper should also have a two page implication of the results section. The paper should follow formal APA style guidelines. Supplemental material on distributions, outliers, influential points, and alternative analyses should be included in an appendix.

Final oral presentation: Students will be asked to make oral presentations of their final reports to the class before submitting the final report, outlining their research question, data, and results (i.e., a short, talk-version of what will be reported in the final paper). This offers the students the opportunity to get feedback before the final paper is turned in. You have around 30 minutes to both present and take questions. I'd suggest speaking for about 20 minutes, leaving 10 minutes for questions.

Textbook: Learning Statistics with R by Danielle Navarro (free PDF download): <https://learningstatisticswithr.com/>

Additional textbooks (optional reading):

Cohen, Cohen, West, Aiken (2003). Applied Multiple Regression/Correlation Analysis Third Edition. New York: Routledge.
Gelman, A., & Hill, J. (2006). Data analysis using regression and multilevel/hierarchical models. Cambridge university press.
Imai, K. (2018). Quantitative social science: an introduction. Princeton University Press.

Administrative Details

Email Policy

We aim to respond to your emails within 48 hours. For this reason, if there are any questions that have some deadlines (e.g., HW or final project), please try to reach out to us at your earliest possible convenience!

How to Upload Homework

You should upload your homework on the relevant spot in "Assignments" on Brightspace before the beginning of the class (2:00PM). Please make sure to submit it as a Word Document, Jupyter Notebook, or R Markdown, with the HW number and your name in the document title (e.g., HW01_Vlasceanu).

Lecture Slides

Slides used in class will be posted to Brightspace after class (typically within a few hours after class).

Policy on Plagiarism

Students are required to abide by New York University's and the College of Arts & Science's academic integrity policies, which can be found here and here, respectively. If you are in doubt about what constitutes academic dishonesty, speak with me before the assignment is due and/or examine the University website. Academic dishonesty includes, but is not limited to, plagiarism of a paper (e.g., taking material from readings without citation, copying another student's paper). Failure to adhere to this policy may result in a failing grade in the class and/or expulsion from the University.

[Policy on Absences](#)

The class will be more valuable and better for everyone if all students show up, ask questions, and contribute to our regularly scheduled meeting times, so it is expected that you attend class and lab. That said, if you do not feel well, please don't come to class, you can log in via zoom, or watch the recording when you feel better. If you are going to miss the exam, for whatever reason, notify me by emailing me before the time of the exam. If you have legitimate health-related or personal problems, you **MUST** make appropriate arrangements **BEFORE** the date of the examination. If you will miss a class, you are responsible for doing the readings, reviewing slides, getting course notes from a classmate to ensure that you are caught up, and turning in assignments on time.

[Policy on Late Assignments](#)

Unless otherwise stated, all assignments are due **AT THE BEGINNING** of class on the date due. Late assignments will not be accepted.

[o Academic Accommodations](#)

Academic accommodations are available for students with documented disabilities. I encourage students with disabilities, including invisible disabilities like chronic diseases or learning disabilities, to contact the Moses Center. Students requesting academic accommodations are advised to reach out to the Moses Center as early as possible in the semester for assistance. I cannot help you if you have not documented things with the Moses Center, so please do use this resource.

Henry and Lucy Moses Center for Students with Disabilities

Telephone: 212-998-4980

Website: www.nyu.edu/csa

Email: mosescsa@nyu.edu