

Professor Ted Gerber

Office Hours: Thursday 12-2, or by appointment
4450 Social Science Building
tgerber@ssc.wisc.edu

TA: Ellie Karls

Office Hours: Wednesday 10-12
Location TBA
lkarls@wisc.edu

TA: Anya Popovych

Office Hours: Thursday 10-12
7102 Social Science
popovych@wisc.edu

INTRODUCTION

This is a course in multiple (or multivariate) regression analysis, a very common statistical technique in the social sciences. Multiple regression models show how a single variable, say “ y ,” is associated with a series of other variables (call them “ x variables”) when their relationships with y are simultaneously controlled. Regression is not only extremely useful by itself; it also serves as a general framework for thinking about relationships among variables. The objective of the course is to familiarize students with the basic design and function of multiple regression models, with how they can be used to test research hypotheses in the social sciences, how to estimate them using real data sets with statistical software, and how to interpret multiple regression results.

By the end of the course, students should have a solid understanding of multiple regression models, including the assumptions on which they are based, the appropriate interpretation of parameter estimates from regressions and inferences about their population values, specific “wrinkles,” such as using dummy variables to specify the effects of one or more categorical variables on y , variations in the effects of one or more x variables depending on the values of other x variables (interactions), and non-linear relationships between one or more x variables and y , how to use a series of regressions to analyze complex causal structures among groups of variables, and how to interpret the results of published research that uses basic linear regression modelling. Students will acquire the ability to perform a wide range of multiple regression analyses using the statistical software package, Stata.

In addition to learning how to use and interpret regression, students will receive instruction in locating and obtaining large, publicly available data sets, defining a research question, setting up a regression analysis to test it, and presenting regression findings in an intuitively appealing and professional fashion. They will also gain the ability to interpret and assess articles in sociological and other social science journals that use regression-based quantitative analysis techniques. The skills covered in the course will provide a foundation for further advanced work in social statistics or for elementary analyses of social science data for research purposes in an academic or professional context.

FORMAT

The course consists of in-person lectures and lab sessions. Lectures will focus on conceptual material, while labs will emphasize software, computer work, and practical tools for completing assignments. Note: all lecture and lab sessions are considered *mandatory*. Missing a few sessions due to illness, travel, or other conflicts is acceptable, and it is not necessary to provide excuses or doctor's notes. Some lectures will be delivered via Zoom (see syllabus for dates). Most or all labs will all involve use of Stata on laptops. You can download a site-licensed copy of the latest version of Stata for free here (use your UW NetID to log into the DoIT site): <https://software.wisc.edu/cgi-bin/ssl/csl.cgi> . If you do not have a laptop, let us know immediately, and we will arrange for you to use one provided from the SSCC's mobile lab. All lab sessions will be run by the teaching assistants, Ellie Karls and Anya Popovych.

BOOK

The sole textbook for the course is: Rachel A. Gordon's *Regression Analysis for the Social Sciences*. Second Edition. Routledge. 2015. *Make sure you have the SECOND edition of the book!* I recommend purchasing the book, as it will likely be helpful to you in the future as a reference. However, you can rent the book online, if you prefer.

REQUIREMENTS

- 1) Students must complete *10 homework assignments*, graded for completeness, timeliness, and accuracy (see grading rubric below), each worth 4 points. Assignments will be posted and submitted on the Canvas website for the course. They must be submitted via Canvas *by 12:00 pm (noon) on Friday* the week they are due. Some assignments will be exercises from the textbook, others will involve special handouts. Homework will include both written problems and statistical software problems. Late homework will be penalized, in fairness to those who submit it on time and in fairness to the TAs. You will receive a grade on homework assignments, but not detailed comments, because the solutions will be covered in lab sessions the Monday following the deadline for the assignments. Overall, **homework is worth 40% of the grade.**
- 2) A two-part *midterm examination* will be held during lecture and lab on October 27 and October 31. The midterm will cover bivariate and multiple regression basics, and also the software skills learned during the first half of the course. The **midterm score is worth 10% of the overall grade.**
- 3) A *take-home final* will be distributed on the last day of class, due on the scheduled exam time for the course, December 20, 3:00 pm. The final will cover material from the whole course, but with a focus on topics after the midterm. The **final is worth 25% of the overall grade.**
- 4) A *seminar paper* presenting an original data analysis. For undergraduates the paper should be 10-15 pages, for graduate students 15-20 pages (double spaced), in length. I will provide a handout with more detailed instructions early in the semester. The paper offers students the chance to execute an original research project using regression analyses. It will be due at 12:00 pm (noon) on December 15. **The paper counts for 25% of the overall grade.**

HOMEWORK GRADING RUBRIC

Each homework assignment is worth 4 points (out of 100) in the calculation of your total grade. Homework grades will always be in 5-point percentage intervals (100%, 95%, 90%, 85%, etc.) They grades will be based on completeness and timeliness (75%) and correctness (25%), as follows:

Completeness and timeliness (75%)

- If all assigned problems are complete and submitted by the due date (and time) you receive the full 75%.
- 5% off for each question that is only partly answered.
- 10% off for each question that is not answered at all.
- 10% off for less than 24 hours late, 20% off for 24-48 hours late, 30% off for over 48 hours late but submitted prior to Monday lab sessions; no credit for assignments not submitted by the following Monday's lab session.

Correctness (25%)

- If all work is completed correctly, you receive the full 25%
- 5% off if most of the work is correct, with a few minor mistakes only
- 10% off if the work is generally correct, but with a few major mistakes or lots of minor mistakes
- 15% off if only about half of the work is correct, multiple major mistakes and/or lots of minor ones
- 20% off if very little of the work is correct
- No credit at all if the work is entirely incorrect

FINAL GRADES

Final grades will be based on the percentage of 100 possible points earned. Regular attendance of lectures and labs is expected: *students who routinely miss either labs or lectures will have up to five points deducted from their average prior to calculating the final grade.* The Teaching Assistants will grade weekly homework assignments, the midterm, and the final exam in a timely fashion, posting grades in Canvas. Professor Gerber will grade the final papers of Sociology PhD students. All other papers will be graded by the TAs. The following intervals will be used to convert point percentages to letter grades:

100 to 92.50:	A
92.49 to 87.50:	AB
87.49 to 82.50:	B
82.49 to 77.50:	BC
77.49 to 70.00:	C
69.99. to 60.00:	D
Under 60.00:	F

ACADEMIC CONDUCT

By virtue of enrollment in Sociology 361, you agree to uphold the high academic standards of the University of Wisconsin-Madison; academic misconduct is behavior that

negatively impacts the integrity of the institution. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these previously listed acts are examples of misconduct which may result in disciplinary action. Examples of disciplinary action include, but is not limited to, failure on the assignment/course, written reprimand, disciplinary probation, suspension, or expulsion. Students and instructors alike are expected to treat one another and with respect and civility.

LEARNING OBJECTIVES

Students completing the course will be ready to conduct their own research using linear regression analysis and secondary data sources, at a level suitable for beginning research assistants on faculty research projects. They will be able to find, download and format social science data suitable for analysis. They will obtain a working knowledge of the Stata statistical software program, including how to use the software to create and transform variables necessary for regression analyses (including scales), how to estimate and refine multiple regressions, and how to implement and interpret tests for the statistical significance of regression coefficients (individually and in combination). They will understand how to make, interpret and critique advanced specifications involving dummy variables, interactions, and nonlinear transformations. They will develop a sense of how theoretical and/or substantive empirical claims can be tested, documented, and supported or refuted quantitatively through regression analysis. They will have the ability to interpret and critique published social science research that uses linear regression. They will be able to write an article-length description of an independent research project that uses linear regression. By the course's end they will have a solid foundation for studying the many methods that are extensions of linear regression, including those that will be covered in Sociology 362.

ACCOMODATIONS FOR STUDENTS WITH DISABILITIES

The University of Wisconsin–Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW–Madison policy (Faculty Document 1071) require that students with disabilities be reasonably accommodated in instruction and campus life. Reasonable accommodations for students with disabilities is a shared faculty and student responsibility. Students are expected to inform faculty of their need for instructional accommodations *by the end of the third week of the semester*, or as soon as possible after a disability has been incurred or recognized. I will work either directly with the student or in coordination with the McBurney Disability Resource Center to identify and provide reasonable instructional accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA.

SCHEDULE

Week 1 (September 7) Introduction

Week 2 (Sep 11/15): Review: Basic concepts, descriptive and inferential statistics

- Sep 11 LAB: Introduction to Stata, data management, and finding datasets on-line
READ Gordon, chs. 1-4
- Sep 15 *Assignment 1 due via Canvas at 12:00 pm*

Week 3 (Sep 18/22): Finish review, t-tests using Stata, start Bivariate Regression

- Sep 18 LAB: Review Assignment 1, Prepare for Assignment 2
READ Gordon, ch. 5
- Sep 21 Lecture delivered via Zoom (do not meet in classroom)
- Sep 22 *Assignment 2 due via Canvas at 12:00 pm*

Week 4 (Sep 25/29): Bivariate Regression and Correlation (cont.)

- Sep 25 LAB: Review Assignment 2, Prepare for Assignment 3
READ Gordon, ch. 5 (again)
- Sep 29 *Assignment 3 due via Canvas at 12:00 pm*

Week 5 (Oct 2/6): From Bivariate to Multivariate Regression

- Oct 2 LAB: Review Assignment 3, Prepare for Assignment 4
READ Gordon, ch. 5 (again)
- Oct 6 *Assignment 4 due via Canvas at 12:00 pm*

Week 6 (Oct 9/13): Multivariate Regression Basics (cont.)

- Oct 9 LAB: Review Assignment 4, Prepare for Assignment 5
READ Gordon, ch. 6
- Oct 13 *Assignment 5 due via Canvas at 12:00 pm*

Week 7 (Oct 16/20): Multivariate Regression Basics (cont.)

- Oct 16 Review Assignment 5, Prepare for Assignment 6
Read Gordon, ch. 6 (again)
- Oct 19 Lecture delivered via Zoom (do not meet in classroom)
- Oct 20 *Assignment 6 due via Canvas at 12:00 pm*

Week 8 (Oct 23/27): Multiple Regression Basics

- Oct 23 LAB: Review Assignment 6, Review for Midterm Exam
READ Gordon, ch. 6 (again)

Oct 26 *Midterm Exam, Part I (in class)*

Week 9 (Oct 30/Nov 3): Dummy Variables

- Oct 30 *LAB: Midterm Exam, Part II (in Lab)*
- Nov 2 READ Gordon, ch. 7

Week 10 (Nov 6/10): Interactions

- Nov 6 LAB: Review midterm exam, Prepare for Assignment 7
 READ Gordon ch. 8
- Nov 10 *Assignment 7 due via Canvas at 12:00 pm*

Week 11 (Nov 13/17): Interactions (cont.)

- Nov 13 LAB: Review Assignment 7, Prepare for Assignment 8
 READ Gordon ch. 8 (again)
- Nov 17 *Assignment 8 due via Canvas at 12:00 pm*

Week 12 (Nov 20/24): Non-linear effects

- Nov 20 LAB: Review Assignment 8, Prepare for Assignment 9
 READ Gordon ch. 9

Nov 23 Thanksgiving, no class!

Week 13 (Nov 27/Dec 1): Finish Non-linear (cont.)/Confounders and Mediators

- Nov 27 LAB: Review interactions and non-linear effects, consult on projects
 READ Gordon ch. 10
- Nov 30 Lecture delivered by Zoom (do not meet in classroom)
- Dec 1 *Assignment 9 due via Canvas at 12:00 pm*

Week 14 (Dec 4/8): Confounders and Mediators (cont.)/Reading regression studies

- Dec 4 LAB: Prepare for Assignment 10, consult on projects
- Dec 7 READ (for discussion in lecture): article by Eng (2012)
- Dec 8 *Assignment 10 due via Canvas at 12:00 pm*

Week 15 (Dec 11/15): Conclusion

- Dec 11 LAB: Review assignment 10, review for final
- Dec 12 Last class: Take home final exam distributed, *due December 20 at noon.*
- Dec 15 *Final papers due at 12:00 pm (noon)!!*