Goals
Students who leave this class should be able to competently use regression analysis to examine substantive questions of interest in the social sciences. The goal of the course is to introduce students to common problems encountered in multivariate regression analysis, examine some solutions to these, and examine extensions of the basic linear model. Extensions include binary and limited dependent variables, complex data structures such as those found in multivariate models, and a preliminary examination of maximum likelihood techniques for regression estimation. The goal in covering these extensions is not to provide a thorough discussion, but instead to expose students to a range of critical techniques in the social sciences.

Meetings
The course will meet once a week in Tarbutton 206. This meeting will be a lecture that focuses on the statistical concepts for the week. For the second scheduled day, we will have an online meeting. You are expected to find a location in which you can access the full version of Stata and the internet in order to complete lab exercises and contact me online. We will conduct these meetings through Adobe Connect, although there is also the possibility of using google chat and Skype to allow easier chatting in person and screen sharing.

In the lab, students will work in Stata. I will help you learn syntax by providing examples, but part of learning to “do” statistics will be investigating syntax on your own and learning how to do the things you need to do. The textbooks will help with this. The UCLA pages on Stata are also excellent resources. One of the other nice things about Stata is that many answers and resources can easily be found online.

Readings
Allison, Paul. D 1999. *Multiple Regression: A Primer*. Pine Forge Press. This is my first time using this book. Students have often requested an easier version of the text – this may be too easy, but we will see how it goes.

Michael N. Mitchell. 2012. *Interpreting and Visualizing Regression Models Using Stata*. Stata Press. This book serves as a more direct introduction to using Stata, particularly some slightly more advanced techniques involving visualization of data. I hope this will serve as a useful set of examples/manual for using Stata.
J. Scott Long and Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. Sage. *This provides a nice discussion of categorical dependent variables and we will rely on it heavily later in the course.*

Should you be interested in other texts for any of these methods, please ask and I can make useful suggestions.

**Homework Assignments**
You will typically have weekly assignments in which you practice skills from the statistical techniques covered in the course. The assignments will be geared toward analyzing data and discussing results, particularly toward application and interpretation. Late assignments will incur a penalty of one point per day late.

**Research Project**
Students will be required to complete a research project. The project will involve analyzing a dataset using the techniques covered in class and presenting the findings in a paper. The goal of the research project is to practice quantitative data analysis in the context of a theoretically-inspired research project. More details on the research project will be provided later.

**Grade Breakdown**
- Engagement/Participation 10%
- Homework assignments 40%
- Mid-term 20%
- Research Project 30%

**Course Schedule**

**January 10 – Regression Review**
Math: Fox (online)

January 12– Intro to Stata through lab.

**January 17 – Bivariate Elaboration & Multivariate Basics (remote)**
Allison, pp. 9-34.
Math: Fox (online), pp. 86-100; 105-112

January 19 – Multivariate regression in Stata, More intro to Stata, ICPSR
January 24 – Categorical Independent Variables
Supplement?? (online); Fox, Chapter 7

January 26 – Categorical data in Stata

January 31 – Nonlinear models
Mitchell, pp. 30-77
Allison, 153-174

February 2 – Visualizing nonlinearity in Stata

February 7 – Interaction effects
Mitchell, 209-246, 275-299, 127-140

February 9 – Interaction effects in Stata

February 14 – Model evaluation and selection
Agresti and Finlay, 441-448

February 16 – Model evaluation in Stata

February 21 – Violations of Model Assumptions, Heteroskedasticity and non-normality
Agresti & Finlay, pp. 283-288, 448-451
Fox, pp. 267-274, 276-277

February 23: lab

February 28 – Outliers and Influential Observations
Online reading

March 2: lab

March 7 – No class (spring break)

March 9 – No class (spring break)

March 14 – Collinearity and Missing Data
Read: Fox, pp. 307-309, 323-325, 548-553, Allison, 137-151

March 16: lab

**March 21 – Binary Dependent Variables and Logistic Regression, Introduction to MLE**
Read: TBD (2nd reading list)

March 23: lab

**March 28 – Binary Dependent Variables and Logistic Regression, Interpreting Coefficients**
Read: TBD (2nd reading list)

March 20: lab

**April 4 – Probit models, Latent variable justification, Understanding Margins, Ordered Logit**
Read: TBD (2nd reading list)

Read: [http://www.academia.edu/2010847/A_Quick_Look_at_the_Margins_Command](http://www.academia.edu/2010847/A_Quick_Look_at_the_Margins_Command), PDF to be posted on blackboard

April 6: lab

**April 11 – Multinomial Logit, Poisson Intro**
Read: TBD (2nd reading list)


April 13: lab

**April 18 – Poisson, Negative Binomial, Zero-Inflation, Tobit/Craggit**
Read: TBD (2nd reading list)

April 20: lab