

---

# POLITICAL DATA ANALYSIS

---

POS 6737 –SECTION 16GF  
DEPARTMENT OF POLITICAL SCIENCE, UNIVERSITY OF FLORIDA  
FALL 2015

**INSTRUCTOR**  
PROF. SUZANNE M. ROBBINS  
205 ANDERSON HALL  
352-273-2381  
[SUZANNE.ROBBINS@UFL.EDU](mailto:SUZANNE.ROBBINS@UFL.EDU)

OFFICE HOURS: WEDNESDAYS 930-1130  
CLASS MEETS: MAT 12, 8:30-11:30, THURS

**GRADER:**  
ALEC S. CHUNG, [ALECCHUNG84@UFL.EDU](mailto:ALECCHUNG84@UFL.EDU)

---

## COURSE DESCRIPTION & OBJECTIVES

---

This course provides an introduction to the theory and practice of quantitative data analysis. Most of the course will focus on probability theory and mathematical statistics and interpretation. The primary objective is to provide the foundation that will be necessary for basic data collection and analysis and for further study in subsequent data analysis courses. At the end of the semester, students should find themselves equipped with the tools to develop their own statistical models for analysis.

The course has three main goals. First, the students are expected to learn how to design and carry out research that employs statistical techniques as a means of testing substantive theories of politics. Second, the students are expected to build a good foundation in statistics that would prepare them for learning more advanced statistical tools and analysis. Third, the students are expected to learn enough statistical skills to be able to understand as well as engage published works in political science research that uses statistical analysis as means of testing theoretical arguments. In the weekly class meeting the course will more or less be conducted as a lecture-based workshop. Labwork will constitute an important component of the learning enterprise – learning how to use available statistical software – Stata– is a must to succeed in this course.

Specific goals this semester include:

- developing testable hypotheses
- collection and manipulation of data
- develop statistical literacy and systematic approaches to research
- summarize and display data accurate and effectively
- compute and interpret descriptive statistics
- construct confidence intervals and test hypotheses for numerical variables (t tests)
- prepare contingency tables and test hypotheses for categorical variables (Chi-sq tests)
- build simple bivariate and multivariate linear regression models and interpret the output
- draw appropriate inferences from the results of statistical analyses and report findings
- interpret the results of research as presented in journal articles and the popular press

- present research findings in written format
- learn basic statistical software

---

## REQUIREMENTS & EVALUATION

---

The requirement for this course is simple: work diligently and persistently. **This includes attending classes**, doing the readings carefully before the seminar meets, and working regularly on the problem sets and the research paper. Each student should expect to be spending many hours learning how to effectively use the statistical Stata 12 (or later) software commonly used to estimate the models discussed in class.

There will be a number of homework assignments that the students must complete and turn in. The homework assignments are due on the specified dates; **no late submissions are accepted**. In addition, students are strongly encouraged to solve the odd-numbered exercises at the end of each chapter of the Agresti textbook. The answers for the Agresti problems are provided at the back of the book. This is a powerful way to put into practice the concepts learned in each chapter as well as provide you with much needed exercise to effectively understand and master the purported statistical skills.

A major component of the course evaluation will be a term research paper. Each student will produce a manuscript of high quality using an appropriate modeling strategy.

Communicating your results to others is as important as getting good results in the first place. Every assignment – homework, exam, paper - requires interpretation and is as important as getting the correct result. Do not submit raw computer output or do files (except for the final paper, in which it is submitted *in addition to* the paper) Assignments should be submitted as a PDF file if possible.

---

## REQUIRED READING MATERIALS

---

Agresti, Alan and Barbara Finlay. 2009. *Statistical Methods for the Social Sciences, 4<sup>th</sup> Edition*. New York: Pearson.

Acock, Alan C. 2014. *A Gentle Introduction to Stata, 4<sup>th</sup> Edition*. College Station, TX: Stata Press.

Additional readings as noted in the course schedule, available on Canvas.

---

## COMPUTER REQUIREMENTS

---

All models in this class can be estimated using the Stata software package. In addition, students should bring laptops to class, as the final hour of each session is devoted to Stata. Stata licenses are available via UF Apps. You may also purchase a personal license directly from Stata (please see me for more information). The best on-line resource for learning Stata is at UCLA:

<http://www.ats.ucla.edu/stat/stata/>.

## DISTRIBUTION OF GRADES

---

### Distribution of Grades

**15%:** Five homework assignments (1-5). Overall the homework assignments will count for 15% of the overall grade. No excuse will be accepted for not turning any assignment (except when justified with officially acceptable documentation). All assignments are due typed and double-spaced at the beginning of class on their respective due dates via Canvas. No late submissions accepted as we will go over the homework in class.

**10%:** Participation. In most seminars, this means raising questions and participating in debates. In this class, attending class is the most critical component of participation. This includes both the lecture session and the lab session. Class and lab time is not a time to be working on material from other classes.

**35%:** Midterm Exam in class. The midterm will consist of problem sets and some definitions. Students will be allowed the use of a calculator and will be provided formula sheets.

**40%:** A research paper on a topic chosen by the student in consultation with the instructor. The goal is to produce a high quality manuscript, using a model (or models) discussed in the course. The research paper consists of five components throughout the semester with the final product due on December 15. More information on the project below.

## OTHER POLICIES

---

Requirements for class attendance are consistent with the attendance policy stated in the Graduate Catalog Regulations found here: <http://gradcatalog.ufl.edu/content.php?catoid=6&navoid=1219>. Attendance is required.

Students with disabilities requesting accommodations should first register with the Disability Resource Center (352-392-8565, [www.dso.ufl.edu/drc/](http://www.dso.ufl.edu/drc/)) by providing appropriate documentation. Once registered, students will receive an accommodation letter which must be presented to the instructor when requesting accommodation. Students with disabilities should follow this procedure as early as possible in the semester.

Information on current UF grading policies for assigning grade points and acceptable graduate-level grades may be found here: <http://gradcatalog.ufl.edu/content.php?catoid=6&navoid=1219>.

Students are expected to provide feedback on the quality of instruction in this course by completing online evaluations at <https://evaluations.ufl.edu>. Evaluations are typically open during the last two or three weeks of the semester, but students will be given specific times when they are open. Summary results of these assessments are available to students at <https://evaluations.ufl.edu/results>.

All work in this class is to be your own. Please take note of the student Honor Code, Student Conduct Code and Standards of Ethical Conduct, which may be found in the Graduate Catalog: <http://gradcatalog.ufl.edu/content.php?catoid=6&navoid=1219>.

---

## RESEARCH PAPER STRATEGY

---

### Each student must:

- Identify a significant research question in his/her field of study
- Choose data from a data set which represents the variables involved
- Conduct analysis to address the research question using one of the techniques discussed in the course.
- The final product should be 15-20 pages long, include statistical analysis, and bibliography.
- Provide replication files (data and do-files).

In order for the instructor to provide guidance in the preparation of this paper, you will be required to turn in various brief intermediate components via canvas throughout the semester. Every student must meet with me a minimum of two times (as noted below) about the project during my office hours.

### DUE DATES FOR PAPER ASSIGNMENTS

---

	<b>date</b>	<b>Assignment</b>
<b>Phase I</b>	10-Sep	Topic of the paper (one paragraph outlining the importance broadly and to the field, plus tentative research question); Students must meet with within a week during office hours to obtain feedback.
<b>Phase II</b>	1-Oct	Research question and hypotheses; Annotated Bibliography, properly formatted, Chicago style
<b>Phase III</b>	5-Nov	Description of the data to be used, how and when it will be obtained; Description of the tentative data analysis technique to be used, how it will be applied, and why it is appropriate. Students must meet with me within a week during office hours to obtain feedback
<b>Phase IV</b>	3-Dec	Data Analysis Results completed to address the question; submit replication files (data files/do files)
<b>Phase V</b>	15-Dec	Complete paper, with title page, intro, lit review, theory, data/methods, analysis/discussion, bibliography and appendices as appropriate. Resubmit replication files (data files/do files).

---

## COURSE SCHEDULE\*

<b>Date</b>	<b>Subject</b>	<b>Reading</b>	<b>Due</b>
<b>27-Aug</b>	Introduction; Data Ethics <i>Getting Started in Stata</i>	Agresti, 1 <i>Acock, 1</i> Rdgs Canvas: King/Sands; News Accounts; Lupia/Elman	
<b>3-Sep</b>	Sampling/Measurement	Agresti, 2	Homework 1
<b>10-Sep</b>	Introduction to Univariate Analysis; Descriptive Statistics <i>Entering Data</i>	Agresti, 3  <i>Acock, 2</i>	Phase I
<b>17-Sep</b>	Probability Distributions; Visualization <i>Preparing Data for Analysis</i>	Agresti, 4 <i>Acock, 3</i>	
<b>24-Sep</b>	Statistical Inference: Estimation <i>Working with Commands, do files, and Results</i>	Agresti, 5 <i>Acock, 4</i> Rdgs Canvas: Gross; Cohen	Homework 2
<b>1-Oct</b>	Statistical Inference: Significance Tests <i>Descriptive Statistics &amp; Graphs</i>	Agresti, 6 <i>Acock, 5</i>	Phase II
<b>8-Oct</b>	Comparison of Two Groups <i>Tests for One or Two Means</i>	Agresti, 7 <i>Acock, 7</i>	Homework 3
<b>15-Oct</b>	Midterm Exam		
<b>22-Oct</b>	Intro to Bivariate Analysis; Associations with Categorical Variables <i>Statistics &amp; Graphs for Two Categorical Variables</i>	Agresti, 8  <i>Acock, 6</i>	
<b>29-Oct</b>	Linear Regression, Correlation; <i>Bivariate Correlation and Regression</i>	Agresti, 9 <i>Acock, 8</i>	Homework 4
<b>5-Nov</b>	Introduction to Multivariate Relations <i>Multiple Regression</i>	Agresti, 10, 11 <i>Acock, 9</i>	Phase III
<b>12-Nov</b>	Comparing Groups, ANOVA <i>ANOVA</i>	Agresti, 12 <i>Acock, 10</i>	
<b>19-Nov</b>	Analysis of Covariance, Dummy Variables <i>Regression Diagnostics</i>	Agresti, 13 <i>Acock, 10</i>	Homework 5 on 11/20
<b>26-Nov</b>	No Class, Thanksgiving		
<b>3-Dec</b>	Model Building, Logistic Regression <i>Multiple Regression, Logistic Regression</i>	Agresti, 14, 15 <i>Acock 10, 11</i>	Phase IV
<b>15-Dec</b>	No class		Phase V

Items in *italics* relate to Stata work.