POLITICAL DATA ANALYSIS

POS 6737 –SECTION 2A48
DEPARTMENT OF POLITICAL SCIENCE, UNIVERSITY OF FLORIDA

INSTRUCTOR
PROF. SUZANNE M. ROBBINS
205 ANDERSON HALL
352-273-2381
SUZANNE.ROBBINS@UFL.EDU
OFFICE HOURS: M/W/F 1-2PM

CLASS MEETS:
WEDNESDAY: 3-4:55PM MAT 251
WEDNESDAY: 5:10-6PM DATA LAB

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, 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. The bulk of learning in the course, however, will take place outside of the classroom. 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)

1

- 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. Each student should expect to be spending many hours learning how to effectively use the statistical Stata 12 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.

REQUIRED READING MATERIALS

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

Acock, Alan C. 2014. *A Gentle Introduction to Stata, 4th Edition.* College Station, EX: Stata Press.

COMPUTER REQUIREMENTS

All models in this class can be estimated using the Stata software package.   You must have a computer account and password to use the computers in the Anderson datalab. Stata is available on all of the Anderson Hall datalab computers. 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/. Stata is also available free for UF students.

DISTRIBUTION OF GRADES

Distribution of Grades

> **25%:** 5 weekly graded homework assignments (1-5): each will count for 5% of the final grade. Overall the homework assignments will count for 25% 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. **No late submission accepted for any**

reason (except when justified with officially acceptable documentation). Homeworks will be submitted via Canvas.

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.

30%: Midterm Exam in class. The midterm will consist of problem sets. Students will be allowed the use of a calculator and will be provided formula sheets.

35%: Take-home Final Exam. Students will analyze a data set as if writing an academic conference paper. The data and minimum components for analysis will be provided in April. The grade will be on the quality of analysis and professionalism of the written presentation. The final exam will be submitted via Canvas.

Many students will find they benefit from extra Stata practice sessions on their own.

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/) 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.

Cell phone use is prohibited during class. Laptops and tablets are discouraged during lecture class sessions (except to access electronic versions of the text).

All work – homework and the final exam - 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. Students who fail to abide by this policy will receive a failing grade on the assignment.

COURSE SCHEDULE

Week	Date	Subject	Reading	Assignment Due
1	7-Jan	Introduction	Agresti, 1	
		Getting Started in Stata	Acock, 1	
2	14-Jan	Sampling/Measurement	Agresti, 2	Homework 1
		Entering Data	Acock, 2	
3	21-Jan	Introduction to Univariate Analysis; Descriptive Statistics	Agresti, 3	
		Preparing Data for Analysis	Acock, 3	
4	28-Jan	Probability Distributions; Visualization	Agresti, 4	
		Working with Commands, do files, and Results	Acock, 4	
5	4-Feb	Statistical Inference: Estimation	Agresti, 5	Homework 2
		Descriptive Statistics & Graphs	Acock, 5	
6	11-Feb	Statistical Inference: Significance Tests	Agresti, 6	
7	18-Feb	Comparison of Two Groups	Agresti, 7	Homework 3
		Tests for One or Two Means	Acock, 7	
8	25-Feb	Midterm Exam		
9	11-Mar	Introduction to Bivariate Analysis; Associations with Categorical Variables	Agresti, 8	
		Statistics & Graphs for Two Categorical Variables	Acock, 6	
10	18-Mar	Linear Regression, Correlation	Agresti, 9	Homework 4
		Bivariate Correlation and Regression	Acock, 8	
11	25-Mar	Introduction to Multivariate Relations	Agresti, 10, 11	
		Multiple Regression	Acock, 10	
12	1-Apr	Comparing Groups, ANOVA	Agresti, 12	
		ANOVA	Acock, 9	
13	8-Apr	Analysis of Covariance	Agresti, 13	
		Multiple Regression	Acock, 10	
14	15-Apr	Model Building with Multiple Regression	Agresti, 14	Homework 5
		Multiple Regression	Acock 10	
15	22-Apr	Logistic Regression	Agresti 15	
		Logistic Regression	Acock, 11	
16	29-Apr	Final Exam due		