

POLS 401: Introduction to Data Analysis
Fall 2007, Rooms BSB 1171
M W 3:00-4:15

Professor Yoram Haftel

Office: BSB 1108 C

Office Phone: (312) 996 8778

Office Hours: M W 1:30-2:30 PM **and by appointment**

Email: yhaftel@uic.edu

Webpage: <http://tigger.uic.edu/~yhaftel/>

Teaching Assistant: Heidi Lawson, hlawso2@uic.edu

Course Description:

Statistics are all around us. You can't open a newspaper, much less a social science journal, without seeing them. Statistical analysis has proven quite useful in helping us understand our world, and is an indispensable scientific tool. However, statistics are quite often abused and misused. This course will provide you with an introduction to applied statistics. We will work on acquiring the basic building blocks necessary to competently consume and use statistics. Our goal here is not only to arm you with knowledge to employ basic statistics in your research and be a discerning critic of the statistics others report but to also give you the foundation that will allow you to become (should you choose) an advanced user of statistics.

Course Requirements:

Our focus will be on the logic of statistical analysis and will mostly be taught at a conceptual level. So there will be minimal reliance on formulas and proofs. I assume that you have done no prior work in statistics. Mathematical knowledge at the level of high school algebra is expected, but, if necessary, we can brush up on those skills. To put it more simply: if you're not mathematically inclined, don't worry; you'll be fine.

Learning statistics is like learning a foreign language. It takes time, repetition, and a lot of practice. Consequently, this will be a hands-on class. You will be asked to complete bi-weekly problem sets. These assignments will give you an opportunity to apply (and even extend) concepts discussed in class. You will be analyzing real data using a computer statistical program (see below). We will devote two classes for a Stata workshop. I expect students to complete their mastery of this or another statistical package outside of class. The ability to navigate new statistical packages with minimal formal help is an important professional skill to have.

I strongly encourage you to work together with each other. Your fellow colleagues are your best resource. There's no better way to learn technical subjects. I do ask that you type up answers to problem sets individually, in your own words. The only exceptions are exams. The midterm and final must be done individually.

Course Evaluation:

Assignments will be weighted in the following way:

Problem Sets	42%
Midterm	30%
Final	<u>30%</u>
Total	102%

The grading scale is as follows: A (88-102); B (76-87); C (64-75); F (<64)

* Political science graduate students who receive C or less must retake this course

Course Policies:

I don't accept late work. Problem sets are due at the beginning of class. If you can't make it to class for some reason, it is your responsibility to get me (or David) the problem set before class. As a general rule I do not give make-up exams. In rare and extreme circumstances I will make an exception, but in order to be considered for an exception, *you must contact me before the exam*. Do not expect sob stories after the fact to move me.

Please attend class. Some people assert that all knowledge is cumulative. Perhaps this is a deep ontological debate in social science, but it isn't in this class. What you learn in earlier classes will be crucial for getting concepts in later classes. You will do yourself a great disservice by missing classes. However, if you find it inevitable to miss a class, make arrangements to get notes from a colleague. Please don't ask me for my notes.

Please feel free to ask questions during class. I know it's a cliché, but it's true that if you are confused about something, others may be as well. Please don't assume you're the only one not getting it. So contribute to the public good and your private gain by asking a question. Also, feel free to come by my office if you need help. The only thing that I ask is that you've thoroughly tried to work through the problem – ideally with classmates – before seeking my help. You'll be surprised how many things you can figure out without my help. If you've tried to understand something and you can't, though, there is no shame in coming to me during office hours (or set an appointment). In fact, I strongly encourage you to talk to me in such instances.

Course Material:

Most of the course readings will come from the following text book, which is available in the UIC bookstore (or online). I know this book isn't cheap, but it offers a perspicuous treatment of beginning statistics – it's not too "mathy" but it doesn't insult your intelligence either. Plus, I think you will find it a useful reference book after the course is over.

Freedman, David, Robert Pisani, and Roger Purves. 2007. *Statistics*. 4th Edition. New York: W.W. Norton.

I will handout in class readings that come from other sources.

Statistics Computing Packages:

In this course we will primarily use Stata, which is widely used by political scientists. A good thing about Stata is that the software is updated every few years with nifty improvements in graphics and command functionality. A bad thing about Stata is the software is updated often, changing command lines. For the purposes of this course, it should not matter whether you are using Stata 7.0, 8.0 or 9.0. We will have a series of short tutorials on how to use Stata. I also strongly recommend that each student owns the following book:

Acock, Alan C. 2006. *A Gentle Introduction to STATA*. College Station: StataCorp.

Academic Integrity:

The university expects students to fulfill their academic obligations through honest and independent effort. Any of the work you do in this course is expected to be your own and according to university and class rules. Any attempt to gain an advantage not given to all students is considered dishonest whether or not it is successful. Academic dishonesty is considered a serious offense subject to strong disciplinary actions, included but are not limited to failing this course. For additional information, see the Student Code of Conduct at:

http://www.vcsa.uic.edu/MainSite/departments/dean_of_students/Links/UIC+Discipline+Code.htm#3

Special Accommodations:

Any Student who has a need for accommodation based on the impact of a disability should contact me privately to discuss the specific situation as soon as possible. Contact the Office of Disability Services to coordinate reasonable accommodations for students with documented disabilities.

Course Calendar

Below is a tentative plan of the course. This will give you a rough idea of the topics we will be covering each week. However, the confidence band around course topics expands the later we get in the course, because I may slow things down (or even speed them up) if necessary.

It is your responsibility to keep track of all dates, assignments, and changes (which will be announced in class).

Date	Topic	Reading	PS
8/27	Introduction and Experiments	Chapter 1	
8/29	Observational Studies	Chapter 2	
9/5	Stata Workshop I		
9/10	Histograms	Chapter 3	
9/12	Stata Workshop II		PS#1
9/17	Summary Statistics	Chapter 4	
9/19	The Normal Curve	Chapter 5	
9/24	Measurement Error	Chapter 6	
9/26	Cross Tabulation	Nachmias & Nachmias: pp. 390-413	
10/1	Correlation	Chapters 8-9	PS#2
10/3	Regression	Chapter 10	
10/8	Root Mean Squared Error	Chapter 11	
10/10	The Regression Line	Chapters 7 & 12	
10/15	Midterm		PS #3
10/17	Probability	Chapter 13	
10/22	Return of Probability	Chapter 14	
10/24	Binomial Formula	Chapter 15	
10/29	Law of Averages	Chapter 16	PS #4
10/31	Expected Values	Chapter 17	
11/5	Normal Approximations	Chapter 18	
11/7	Sampling	Chapter 19	
11/12	Chance and Sampling	Chapter 20	PS #5
11/14	Confidence Intervals	Chapter 21	
11/19	Uncertainty and Averages	Chapter 23	
11/26	Measurement Error	Chapter 24	PS #6
11/28	Tests of Significance	Chapters 26 & 27	
12/3	Inference and Cross-Tabs	Chapter 28	
12/5	Wrap Up and Review		PS #7
12/14	Final Exam (3:30 – 5:30)		