Instructor: Meredith Phillips  
Class Meeting: Public Affairs 2343, Mon. & Wed., 9:30-10:50 a.m.  
Prof. Office Hours: Monday, 12:00-3:00 p.m. (Sign up on office door.)  
Office Location: 6323 Public Affairs  
Phone: 310-794-5475  
Email: Meredith.phillips@ucla.edu (please write PP203 in the subject line)  
TA: Tom Jacobson  
TA Email: thomasjacobson@ucla.edu (Please write PP203 in the subject line)  
TA Office Hours: Tues., 5:00-6:00 p.m. and by appointment, TA office, Public Affairs 6243  
Lab Days & Times: We have two labs scheduled in the Luskin computer lab (Wed., 6:00-7:00 p.m. and Thurs., 3:30-4:30 p.m.). You should plan to attend one of these labs (we’ll distribute a signup sheet the first day of class). The lab has also been reserved for an additional hour after the official lab ends so that you can stay in the lab to work on your problem set if you’d like.

Course Description: This course introduces students to basic statistical concepts using real-world examples and hands-on data manipulation. We cover displaying and summarizing data; simple probability and conditional probability; estimating population values from samples; hypothesis testing (including t, F, and chi-squared tests); correlation; simple regression; and multiple regression.

Goals: The primary goal of this course is to help you develop statistical skills that you can use throughout your professional and civic life. After taking this course, you should be able to:
1) interpret and think critically about statistical data discussed in reports and newspapers;
2) conduct basic statistical analyses that are appropriate to the research or policy questions you want to answer; and
3) communicate quantitative results to lay audiences.

Textbook: The perfect book for teaching statistics to public policy students does not exist. The textbook that we will use exclusively until we get to simple regression is:


Although this book is a little weak on public policy and social science examples, it is thoroughly readable—clear, witty, informal yet still informative. Of all the statistics texts I’ve read, I’ve concluded that this text is the least likely to put you to sleep.

Make sure you read the text inside the green boxes as well as the “Where are we going?” introduction to each chapter.
I’m not thrilled with the order in which the book covers topics, so we’ll skip around a lot. Make sure to **pay close attention to the syllabus so that you know what to read for each class period**.

When we start learning about regression, I will also ask you to read:


This book offers an excellent, applied introduction to data analysis and interpretation. You’ll continue to use this book in the second quarter of our statistics sequence.

**Data:** We will use myriad examples from three national data sources. The General Social Survey (GSS) is an almost annual (biennial as of 1994) personal interview survey of U.S. households conducted by the National Opinion Research Center (NORC). The subset of variables I have extracted for this class emphasizes data from 2012, though it also includes several time series that extend back to the early 1970s. I have chosen variables that reflect a wide range of substantive policy interests, including government, voting behavior, race and ethnicity, poverty, the environment, drugs and crime, education, transportation, gender, immigration, and health care. For information about sampling, weighting, question wordings, etc., look here: [http://www3.norc.org/GSS+Website/Documentation/](http://www3.norc.org/GSS+Website/Documentation/).

In lecture, and in some problem sets, we will also use several extracts from the public-use versions of the Early Childhood Longitudinal Survey of Kindergarteners (ECLS-K) and the Education Longitudinal Study of 2002 (ELS). The ECLS-K surveyed a nationally representative cohort of kindergartners in 1998 and has followed them through the end of eighth grade. The ELS surveyed a nationally representative cohort of tenth graders in 2002 and resurveyed them in 2004 and 2006. Both ECLS-K and ELS collected data not only from the students themselves, but also from their parents, teachers, and schools. The questionnaires used to collect the data are available for download at [http://nces.ed.gov/ecls/kinderinstruments.asp](http://nces.ed.gov/ecls/kinderinstruments.asp) and [http://nces.ed.gov/surveys/els2002/questionnaires.asp](http://nces.ed.gov/surveys/els2002/questionnaires.asp).

In addition to using these data for your problem sets, I encourage you to “play around” with the data for fun and see what you discover. The data sets will be available on the network for use in the lab. They will also be available for download from the course web page.

We will construct a fourth dataset from a survey filled out by YOU. We will analyze these data both by hand and on the computer.

**Statistical Software:** We will use Stata to analyze the data in this class. Stata is currently the statistical software of choice in public policy, economics, and sociology. It also resembles other popular packages, like SAS, SPSS, and R. I recommend learning the basics of these other packages, especially R, during your time at UCLA. Statistical computing services at UCLA is a wonderful resource for learning how to use these other software packages (see [http://www.ats.ucla.edu/stat/seminars/default.htm](http://www.ats.ucla.edu/stat/seminars/default.htm)).
Stata is on the computers in the Public Affairs lab and is also available to use remotely. You may want to purchase your own copy of Stata for a substantial student discount (see http://www.stata.com/order/new/edu/gradplans/gp-direct.html). If you decide you want to own a copy of Stata, I would recommend buying the latest version (v. 13) and a perpetual license (the student cost is $189) so that you can use the software beyond this year. But you can also buy a one-year license for $98 or a six-month license for $69. If you decide to purchase Stata, buy either the Intercooled or SE version. **DO NOT buy the Small version.** Note that you do NOT have to buy Stata for this class in order to do your problem sets from a remote location.

If you have questions about your Stata analyses, Academic Technology Services (ATS) at UCLA provides a very helpful website for Stata users (http://www.ats.ucla.edu/stat/stata/default.htm). Whenever you are unsure how to get Stata to do what you need it to do, the ATS website should help. You can also check out other sites that have useful Stata help, including: http://www.Stata.com/support/faqs/ and http://www.cpc.unc.edu/research/tools/data_analysis/statatutorial/.

**Calculators:** You should own a calculator that has statistical functions that calculate the mean and standard deviation. You can purchase such calculators for $10-$20 at many drug stores or at the campus store. If your calculator does not have a button that looks like a summation sign with a plus/minus following it, you probably need a new calculator. If in doubt, ask me after class. **DO NOT WAIT UNTIL THE NIGHT BEFORE THE MIDTERM TO LEARN HOW TO USE YOUR CALCULATOR!**

**Requirements:**

**Reading:** You should do the readings listed for each class **BEFORE** that class. The vast majority of the readings are in the textbook. Other readings will either be handed out in class or will be available on the course website for you to download and print. I will NOT cover everything in the readings in class, so it is imperative that you keep up with the reading.

**Notes:** Before most classes, I will post some very basic notes on the web. You should print those notes and bring them to class so that you can write your own notes directly on the print-outs (or, if you prefer, bring your laptop/tablet to class and annotate the notes electronically). If I have not posted notes by 6:00 p.m. the evening before class, I will distribute copies of the notes in class.

**Class participation:** Class attendance is required. Please arrive **on time.** Some of the material covered in class will not be available in the text. If you are confused about something we discuss in class, **INTERRUPT me and ask.** I also encourage you to ask questions during student presentations. Class attendance and participation count for 10% of your grade.

**Weekly lab sessions:** Most weeks you will have a problem set that requires Stata. In the Public Affairs lab, the TA will provide a quick overview of the problem set and then be available to troubleshoot as problems/questions arise. **Do not expect that you will be able to complete your problem set during the lab session.** Typically, you will need to spend considerable time working on it at a later time. After an hour in the lab, the TA will begin to review material that students
struggled with on the last problem set. During this time you may either stay in the lab to work on your problem set, ask the TA for help, or go home.

**Problem Sets:** Weekly problem sets, assigned at the beginning of class most Wednesdays and due at the beginning of class the following Wednesday, will give you a chance to practice what you’ve learned. Some of the problems will require hand calculations. Others will require computer calculations using Stata. Many will require graphing in Excel or Stata. The problem sets are required and will be graded on an excellent/satisfactory/unsatisfactory basis. We will post solutions on the course webpage. **Make sure you download these solutions and compare them to your own problem set so that you can assess and improve your understanding of the material.** I encourage you to work with others on the problem sets, but make sure you submit your own work (and understand what you are doing!!). You must keep up with the problem sets in order to be prepared for the exams. The problem sets will constitute 15% of your grade.

**Projects:** Each class member will complete one project during the quarter. You will sign up for projects the first week of class. Working in teams, you will collect some data, write a short report, and present your findings to the class. You will be held to strict time limits during the presentations. You will be graded on the quality of both the presentation and your written report. I will give you a set of “writing rules” to follow when editing your report. At the end of the quarter, I will ask each group member to rate the contributions of the other members to the project and will take these ratings into account when assigning project grades. The project will constitute 15% of your grade.

**Mid-Term Exam:** This exam will assess your understanding of material covered during the first half of the course. You may bring into the exam a 4-by-6 note card covered front and back with any information you want. The mid-term will constitute 25% of your grade. I will post last year’s mid-term on the web so that you can get a sense of the types of questions to expect.

**Final Exam:** This exam will be cumulative (i.e., it will assess your understanding of material covered during all ten weeks of the course). As with the mid-term, you may bring in a 4-by-6 note card covered back and front with formulae, memory joggers, and so on. The final will count for 35% of your grade. I will post last year’s final on the web so that you can get a sense of the types of questions to expect.

**Accommodations:** If you wish to request an accommodation due to a disability, please contact the Office for Students with Disabilities as soon as possible at A255 Murphy Hall, (310) 825-1501, (310) 206-6083 (telephone device for the deaf). Website: [www.osd.ucla.edu](http://www.osd.ucla.edu).

**Academic Dishonesty:** Please hold yourself to high standards of academic honesty in this class and always. Please familiarize yourself with UCLA rules relating to academic conduct because I will not hesitate to report any infractions of the rules. What counts as dishonesty in this class, specifically? **Problem Sets:** Submitting a problem set that contains results or calculations that you had no role in producing; submitting a problem set that uses someone else’s words rather than your own to describe or interpret the results; submitting the same problem set as someone else in your class or a previous class.
Projects: Claiming authorial credit on a project to which you contributed little; submitting a project based on data collected by students in a previous class; submitting a project written by students in a previous class; submitting a project that does not attribute information, data, prose, or ideas to their source.

Exams: Using any source of information during the exam other than your brain, your 4-by-6 note card, or me; changing your answers on an exam after it has been returned to you in order to claim that it was mis-graded.

Re-grades: We will not re-grade problem sets. If you feel that your midterm or final has been mis-graded, you may request a re-grade in writing by re-submitting the exam with an explanation of which questions you feel have been graded incorrectly. Please be aware that if you request a re-grade, we will re-grade the entire exam, which means that your score may, in fact, go down. If the points on your exam have been added incorrectly, please simply speak to the TA (you needn’t submit a written re-grade request in that case).

Grades: We ask that you please identify yourself by your university ID number (instead of your name) on all problem sets and exams. Generally, a “B+” indicates good understanding of all course material; a “B” indicates adequate understanding; and an “A” indicates complete mastery of the material. A grade of B- or below indicates that you have not mastered a considerable portion of the course material. I do not assume that grades will be normally distributed. In other words, it is possible for everyone to earn A’s and also possible for everyone to earn B’s or C’s.

COURSE SCHEDULE

Week 1
Meeting 1, Mon, 10/6 – Course Overview; Introduction to the Data; Project Groups; Course Introductions

Meeting 2, Wed, 10/8 – Research Designs; Sampling Basics; Displaying Data
Read for Today:
• De Veaux et al., Chapters 1-3; Chapters 12-13

Lab—Fill out online survey while in lab. Open class data sets. Tabulate variables.

Week 2
Meeting 3, Mon, 10/13 – Measuring the Center and Spread of Distributions
Read before Today:
• De Veaux et al., Chapters 4-5

Meeting 4, Wed, 10/15 – Simple and Conditional Probability; Expected Value
Read for Today:
• De Veaux et al., Chapters 14-16

Available for Download:
Problem Set #1
Projects for Groups 1 and 2
Lab – Making and Describing Data (Problem Set #1)

Week 3
Meeting 5, Mon, 10/20 – Probability Models
Read for Today:
- De Veaux and Velleman, Chapters 6 & 17

Meeting 6, Wed, 10/22 – Sampling Distributions
Read for Today:
- De Veaux et al., Chapter 18
Due: Problem Set #1
Presentations: Groups 1 and 2
Available for Download:
    Answers to Problem Set #1
    Problem Set #2

Lab – Meet in the lab if you are struggling with Stata or had trouble with the first problem set. Otherwise, use the lab time to work on probability (Problem Set #2) on your own or with your study group.

Week 4
Meeting 7, Mon, 10/27 – Standard Errors/Confidence Intervals for Means and Proportions
Read for Today:
- De Veaux et al., Chapters 19 & 23 (stop when you get to the part on hypothesis tests)

Meeting 8, Wed, 10/29 – Hypothesis Testing: One Proportion or Mean
- De Veaux et al., Chapters 20 & 21, and the rest of Chapter 23.
Due: Problem Set #2
Available for Download:
    Answers to Problem Set #2
    Problem Set #3
    Projects for Groups 3 and 4

Lab – Sampling Distribution Simulation, Standard Errors and Confidence Intervals, Hypothesis Testing (Problem Set #3).

Week 5
Meeting 9, Mon, 11/3 – Hypothesis Testing: Power, Comparing Two Proportions or Means, Paired T-tests
Read for Today
- De Veaux et al., Chapters 22, 24, 25

Meeting 10, Wed, 11/5 -- Hypothesis Testing: Comparing more than two means (ANOVA)
Read for Today:
- De Veaux et al., Chapter 28
Presentations: Groups 3 and 4  
Due: Problem Set #3  
Available for Download:  
Answers to Problem Set #3

Lab – Review for Midterm (come with questions!)

Week 6
Meeting 11, Mon, 11/10— Hypothesis Testing: Two Categorical Variables  
Read for Today:  
De Veaux et al., Chapter 26

Meeting 12, Wed, 11/12 – Midterm  
• Bring your calculator and a 4-by-6 card covered front and back with anything

Available for Download:  
Problem Set #4  
Projects for Groups 5 and 6

Lab – T-tests, ANOVAs (Problem Set #4).

Week 7
Meeting 13, Mon, 11/17 Correlation and Simple Regression  
Read for Today:  
• De Veaux et al., Chapters 7, 8, 9, 10

Meeting 14, Wed, 11/19 – Simple Regression  
Read for Today:  
• De Veaux et al., Chapter 27 (stop when you get to logistic regression)  
• Wooldridge, Chapters 1 & 2  
Presentations: Groups 5 and 6  
Due: Problem Set #4  
Available for Download:  
Answers to Problem Set #4  
Projects for Groups 7 and 8  
Problem Set #5

Lab – Chi-Square, Correlation, Simple Regression (Problem Set #5).

Week 8
Meeting 15, Mon, 11/24 – Regression with Dummy Variables vs. ANOVA; Multiple regression

Read for Today:  
• De Veaux et al., Chapter 30  
• Wooldridge, Chapter 7 (just through 7.3)
Meeting 16, Wed, 11/26 -- Class cancelled for Thanksgiving… use the extra time to CATCH UP ON THE READING if you’ve fallen behind. Eat lots of pie to sustain you.

No lab—Happy Thanksgiving!

Week 9
Meeting 17, Mon, 12/1— Multiple Regression
Read for Today:
  - Wooldridge, Chapter 3
  - De Veaux et al., Chapter 31

Meeting 18, Wed, 12/3— More Multiple Regression
Read for Today:
  - Wooldridge, Chapter 4

Presentations: Groups 7 and 8
Due: Problem Set #5
Available for Download:
  - Answers to Problem Set #5
  - Projects for Groups 9 and 10
  - Problem Set #6

Lab – Multiple Regression (Problem Set #6).

Week 10
Meeting 19, Mon, 12/8— More Multiple Regression and Interpreting Tables in Real Studies
Read for Today:
  - TBA (studies will be posted on course website)

Meeting 20, Wed, 12/10— More Interpreting Tables in Real Studies/Summary
Presentations: Groups 9 and 10
Due: Problem Set #6 (won’t be graded but will be marked complete/incomplete)

Lab— Review problem set #6; review for final (come with questions!)

Final Exam
Cumulative Final Exam: Tuesday, December 16, 3-7 pm, Public Affairs 2343