Political Analysis (PLS 300-01) Fall 2013

Course Prerequisite: STA 215 (Introductory Applied Statistics) or equivalent credit
Class Location and Time: Tuesdays in ASH 2120 (a regular classroom) and Thursdays in HRY 114 (A computer lab classroom), 10:00 – 11:15 a.m.
Professor: H. Whitt Kilburn, Ph.D.
Office: 1103 AuSable Hall
Office Telephone: (616) 331-8831
Email: kilburnw@mail.gvsu.edu
In person Office Hours: Tuesday and Thursday, 9-10 a.m., 1-2 p.m., and by appointment. Virtual hours — anytime — when I'm logged onto GVSU Gmail (mail.gvsu.edu) chat.
Course Websites: There are three!

1. The main website, including everything related to data analysis assignments and class handouts: `http://faculty.gvsu.edu/kilburnw/PLS300.html`.
2. We will use a web service, *Nota Bene*, hosted by MIT, for collaboratively reading and discussing various PDF documents, in particular one of the course textbooks. You will receive an invitation to join this site. Save the password and bookmark it.
3. You'll post your data analysis work on Blackboard, through `http://mybb.gvsu.edu`.

This course provides an introduction to fundamental concepts in data analysis and quantitative research methodology for political science. There is much to be learned about data analysis by actually doing it, rather than reading about it. But first, rather than learn by trial and error, we will identify some principles that should inform quantitative political science research at its inception. The overarching goal of the course will be for us to learn how to manage common sources of data in political science and apply fundamental tools of statistical modeling to illuminate the political world.

Good data analysis is not a mechanical, purely objective process. It requires subjective judgment, sound theory, and testable hypotheses, all of which become better with practice — and familiarity with 'best' practices. So we will read what others have written on the subject of quantitative research methodology and will focus our attention on developing fundamental skills in four main areas: 1) research design, 2) theory development and hypothesis testing, 3) data organization and analysis, and 4) research writing and oral communication. The third area will require you to draw upon what you have learned in Statistics 215. Assignments will include a semester long original research paper, numerous in-class and homework problems, an oral presentation, and in-class examinations.

Course texts

Agresti, R. and B. Finlay. 1997. *Statistical Methods for the Social Sciences.* Third Edition. New York: Prentice Hall.

The text covers introductory to intermediate applications of statistics in the social sciences. Less than half of the text covers subjects you studied in Statistics 215 or its equivalent. We will review these concepts and consider more advanced applications of each in political science. In the second half of the course, we will study multi-variate modeling techniques, all of which have important applications in the study of politics. Each chapter includes many

practice problems and answers to odd-numbered exercises. Some problems will be assigned as homework or an in-class quiz.

Tufte, E. 1974. *Data Analysis for Politics and Public Policy.* Englewood Cliffs, NJ: Prentice Hall.

Perhaps more so than any other political scientist, Tufte has influenced how many people think about data analysis. We will read Tufte's book to learn and apply principles of good research design to our own work. After joining *Nota Bene*, access a PDF copy of the book here `http://nb.mit.edu/f/7999`.

Meys, J. and A. de Vries. 2012. *R for Dummies.* (2nd edition). New York: For Dummies.

Meys and de Vries' book is an introduction to the software we will use throughout the course. Since it is written in the accessible "for Dummies" style, you will be expected to read through the chapters on your own. From time to time, I will point out particular sections of the book that are useful for completing the data analysis assignments.

scholarly journal and other articles: Articles from scholarly journals and other sources will be assigned throughout the semester. When possible, PDF copies will be posted on *Nota Bene*. For others, a simple search on the library homepage should bring up each article.

other suggested reading: A series of books on the visual presentation of information are available from the Pew library: Nancy Duarte, *Slideology: The Art and Science of Creating Great Presentations*; and Edward Tufte, *The Cognitive Style of PowerPoint* and *The Visual Display of Quantitative Information.* These references are excellent resources for preparing presentations, developing analysis and graphics. If interested in getting serious about data analysis and communicating the results to others, I suggest taking a look at each one.

statistical data analysis and visualization software We will spend some time reviewing fundamental concepts of statistical inference through paper and pencil exercises. And we will use some online data analysis resources. But for most analyses of data, we will use a software application called R, which is a freely available application for managing, analyzing, and displaying quantitative data. It is widely regarded as the emerging *lingua franca* of data analysis.[1] And it is very popularly used in political science. Given the sea of information in which we live our lives, being skilled in making sense of data is an essential liberal arts skill. Our use of R and other resources will help you to learn these skills and to interpret the work of others. On multiple occasions, I have received emails and phone calls from slightly anxious GVSU students on an internship, looking for help with their assignment to "do something with the data". So let's learn these skills now. Nevertheless, the learning curve is steep. So you will need to be patient and persistent. But the rewards for doing so are great. In the computer labs, R is located under "Departmental Applications" folder "Statistics".

R tutorial website: While optional, some of you may find it easier to get started with the introductory lessons on R at 'code school', `http://www.codeschool.com/courses/try-r`. You can work at the lessons on your own pace, but doing so is optional.

[1] Ashlee Vance, "Data Analysts Captivated by R's Power ", *New York Times*, January 7, 2009, page B6.

Schedule and Activities

Tuesdays and Thursdays On Tuesdays, we will meet in the regular classroom to discuss applications of statistical modeling in political science, and more general principles of research methodology. For Tuesday classes, typically a series of reading assignments will be due, both from the Agresti and Finlay textbook and either Tufte or journal articles. There will be occasional paper and pencil data analysis assignments on Tuesdays. Thursdays will be spent in the computer classroom to refine our skills at the analysis of datasets. Data analysis assignments, either based on paper-and-pencil, or in use of statistical software, will be assigned these days.

Data Management and Analysis Each week of class will involve a homework assignment, either doing data analysis problems, reflecting on concepts from Agresti and Finlay, or a combination of both. Some assignments will involve refining data management skills. Approximately 12 data analysis assignments will be distributed throughout the semester.

These assignments will completed and submitted as an online journal of data analysis, posted and updated throughout the semester. Each student will have a journal on Blackboard, each one freely viewable to other students. You are all welcome to work collaboratively on the data analysis assignments, but only in the sense that you decide collectively how to best approach an answer to each question, exchange tips, and troubleshoot problems. Each student, however, must post his or her own work to their own individual journal. So you can not simply copy and paste someone else's R code or graphics to your own. (Beyond a poor strategy for learning, doing so would be an honor code violation.)

Reading Questions For some Tuesday class meetings, you will be expected to come to class having answered a series of questions on the assigned readings. Your answers must be *type-written* in hard copy, ready to be submitted at the beginning of class.

Research Paper You will be expected to write a research paper on an assigned topic. The project will involve defining a research problem, conducting a literature review, outlining a theory, collecting data, writing hypotheses, and testing the results using the quantitative data analysis skills we will develop. This research paper is required of all students. You will present your research paper to the class at the end of the semester. Papers will be at least 15 pages (double spaced), with additional pages for the presentation of data analyses, any appendices, and bibliography.

Your research paper will make use of at least one nationally representative survey dataset, from within the U.S.A. or abroad, thus giving your the opportunity to explore research questions drawn from domestic politics or the domestic politics of another country.

Research Critique You will be expected to write one summary memo critiquing the application of research methods and statistical modeling we study in class. The subjects of the research critiques appear in the course schedule; some have a close correspondence to the course schedule, others less so. Pick a subject that interests you.

When turning it in, follow these guidelines: The research design critique is intended for you to apply your knowledge and skills evaluating social/political research to a professional application. The assignment is not meant to be a major paper writing assignment, although you should plan to write approximately (no more than) 4 double-spaced pages.

In your memo, you should develop a thesis — your perspective on the research. And a good approach is to make sure you include most of the items below:

1. an identification of the research question
2. a statement of the key hypotheses
3. a discussion of the key theoretical claims
4. a discussion of how major theoretical concepts are empirically measured
5. a description of how data are collected, and causal claims established, in a way that provides empirical evidence bearing on the theory, and in particular whether the study successfully controls for problems of 'spuriousness'.
6. A discussion of what the statistical tests are intended to accomplish and what is found in the results.
7. a summary of the overall results of the studies, and a discussion of the importance of the results.

Examinations A midsemester examination will cover ideas in research design and data analysis from the first half of the course. The equivalent of a final examination will be broken up into small parts and scheduled as in-class short examinations throughout the second half of the course.

in-class and homework activities There will be frequent in-class and homework exercises. The exercises will be announced in class. To receive credit for each one, you must be prepared to turn in the assignment when collected. Late assignments may be submitted for half credit (50%), up to one week past the due date. Work not completed at all receives a 0%.

General Policies and Procedures

I will hold you accountable to high standards. Beyond that, there are a few key things you should know about my expectations of you:

1. Most important of all, while I do not take attendance, I do keep track of tardiness and participation. I expect you to arrive to class on time and actively participate. The class suffers when any student routinely skips class or is apathetic about assignments. If you are in class, I expect you to *be in class*. So please do not sit in your desk, preoccupied with your phone. The same policy applies to laptops; avoid using your laptop to "take notes". But if you insist on watching cat videos throughout class (as in, "But I'm just taking notes!"), don't expect any sympathy from me when you encounter difficulties in the course.

2. Assignments should be turned in by the deadline. There are no make-ups for any work assigned and completed in-class, except for the exams, which must be arranged ahead of time. I do not accept assignments via email. Assignments must be turned in via hard copy or Blackboard, as directed.

3. You should *always* bring relevant materials with you to class, whether the assigned reading, a calculator, or responses to the homework. Be prepared.

4. If you believe you will have trouble with the course, I encourage you to call an academic counselor at the Advising Resources and Special Programs Unit, 331-3588.

Grades

Periodic Grading of Data Analysis Journals 20%
In-class quizzes, other assignments 25%
Research Design Critique 10%
Mid-Semester Examination 20%
Research Paper and Presentation 25%

Unless standardized scores are necessary, the grading scale will be:

A 93-100	A- 90-92
B+ 87-89	B 84-86
B- 80-83	C+ 76-79
C 72-75	C- 70-72
D+ 65-69	D 60-64
F <60	

Code of Academic Integrity: I expect everyone to abide by the GVSU code of conduct regarding academic honesty. You should consult the student handbook and GVSU student bulletin for relevant policies about academic honesty.

Accommodations: If there is any student in this class who has special needs because of a learning, physical, or other disability, please contact the Disabilities Support Services (DSS) Program at 331-3588.

HOW TO SUCCEED IN THIS COURSE

1. Studiously avoid last-minute preparation and complete all assignments.
2. Attend every class session.
3. Participate regularly in class discussions.
4. Talk to me during office hours about any concerns or to clarify expectations.

Schedule

Progress throughout the course readings depends partially on data analysis assignment completion, class comprehension of concepts, and other factors due to our split schedule between computer lab and regular classroom. Changes to the syllabus should be expected!

Data analysis labs are distributed on the day appearing next to each assignment, and are due within one week.

All URLs are active, clickable links in the PDF syllabus. You should complete the reading assignments by the date immediately next to or above each chapter or article.

1 Course Overview, Brief Quiz, and Introduction to R

Tuesday, 8/27: Quiz and introduction to course

Thursday, 8/29: Computer lab meeting today. *Data analysis assignment distributed in class.* Arrive on time, or else you will fall behind and it will be very difficult to catch up.

In preparation for Thursday's class, read the chapters 1 and 2 of *R for Dummies*. Try downloading and installing R on your own personal computer. Work through some simple arithmetic with the command line and any other examples.

We are not meeting next Tuesday. Look ahead to the reading assignments for next week.

2 Univariate and Bivariate Data Visualizations

Tuesday, 9/3 : Labor Day holiday, no class scheduled.

Thursday, 9/5 : Read Chapters 1 "Introduction", 2 "Sampling and Measurement", and 3 "Descriptive Statistics" of Agresti and Finlay. *Data analysis assignment distributed in class.*

We willl learn to use the *lattice* graphics package in R. *lattice* is a system for displaying the distribution of a variable, or the relationship between two variables, across levels of a third. While optional in preparation for the class, Chapter 17 of *R for Dummies* covers the *lattice* package. Review the chapter when completing the data analysis assignment.

If you need additional help beyond class materials, read a short overview of common uses of the *lattice* package: `http://www.bioconductor.org/help/course-materials/2008/advanced_R/latticeLab.pdf`. A more advanced overview of the lattice package by one of the authors of R, `https://www.stat.auckland.ac.nz/~ihaka/787/lectures-trellis.pdf`.

For additional applications of Cleveland's 'Dot Plots' used in the first data analysis assignment, one plot type in *lattice*, See Jacoby, William. Spring 2006. "The Dot Plot: A Graphical Display for Labeled Quantitative Values". The Political Methodologist Newsletter of the Political Methodology Section, American Political Science Association. See page 6, `http://nb.mit.edu/f/8065`.

3 Review of Probability and Uncertainty in Political Science Research

Tuesday, 9/10 : Read Chapter 4, "Probability Distributions", Agresti and Finlay. Read Chapter 1, "Introduction to Data Analysis", Tufte.

Option for a Research Design Critique. Why do violations of human rights occur? What political, economic, and military factors might explain their occurrence? Poe, Steven C. and C. Neal Tate. 1994. "Repression of Human Rights to Personal Integrity in the 1980s: A Global Analysis." *The American Political Science Review.* 88(4):853-872. Due within two weeks of this date.

Thursday, 9/12 *Data analysis assignment distributed in class.* Read Chapters 3-4 of *R for Dummies.* Additional pages announced in class.

4 Estimation, Inference, and Problems in Research Design

Tuesday, 9/17 : Read Chapter 5, "Statistical Inference: Estimation". Read Chapter 2, "Prediction and Projection", Tufte.

Option for a Research Design Critique How does the experience of forced participation as a child soldier in Uganda affect an adult's civic and political engagement? Blattmann, Christopher. 2009. "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review* 103(May):231-247. Due within two weeks of this date.

Thursday, 9/19 : *Data analysis assignment distributed in class.*

5 Significance Testing, Applications of Inference

Tuesday, 9/24 : Read Chapter 6, "Significance Tests" and 7, "Comparison of Two Groups", Agresti and Finlay.

Thursday 9/26 : *Data analysis assignment distributed in class.*

6 Bivariate Association

Tuesday, 10/1 : Read Chapter 8, "Analyzing Association Between Categorical Variables", Agresti and Finlay.

Option for a Research Design Critique: The validity of much research on the effectiveness of charter schools, compared to their 'regular' counterparts, hinges on the ability of researchers to assess educational outcomes across the two types of schools, while controlling for student background and ability as they enter the schools. One of the more rigorous attempts to control for differences, from a researcher at Stanford's Hoover Institution, found that only about 17 percent of charter schools from a pooled national sample produced better outcomes. `http://credo.stanford.edu/reports/MULTIPLE_CHOICE_CREDO.pdf` . Due within two weeks of this date.

Thursday, 10/3 : *Data analysis assignment distributed in class*

7 Modern Opinion Polling and Analysis; Midsemester Examination

Tuesday, 10/8 : Browse the American Association for Public Opinion Research (AAPOR) 'FAQS' on modern survey research at `http://www.aapor.org/Poll_andamp_Survey_FAQ/5511.htm`. Read the AAPOR task force report on non-probability methods of opinion polling, executive summary `http://nb.mit.edu/f/8404`. Within the full report, read Chapter 3, "Introduction to Non-Probability Sampling" at `http://nb.mit.edu/f/8405`.

Volunteering to answer surveys as part of the 'panel' at `https://today.yougov.com/opi/` joins you in a non-probability survey. Check out the surveys and results at YouGov in preparation for today's class.

Option for a Research Design Critique: When good polls go bad. The most recent Minnesota Governor's race was close enough for an official recount. Yet one pre-election poll by a

political scientist and psychologist at the University of Minnesota publicly called the race for Dayton (the Democrat), estimating that he was way ahead of the Republican. Their experience is a cautionary tale of what can go wrong in contemporary election polling, even among well-funded, knowledgeable professionals. See a description of their poll `http://www.minnpost.com/braublog/2010/12/17/24352/mprhumphrey_institute_poll_review_too_many_612s` and the results of two reviews, an internal review `http://minnesota.publicradio.org/documents/news/2010/12/HHH_MPR_Internal_Report_of_Polls_2010_Final.pdf`, and an external review of the poll by a director of the Gallup poll and an ethics compliance official for the American Association for Public Opinion Research `http://minnesota.publicradio.org/documents/news/2010/12/MPR-HHH_Polls_2010_Newport_Review.pdf`. The memo should be posted to Course Documents within two weeks of this date.

Thursday, 10/10 : **Midsemester Examination in class.**

8 Simple to Multiple Linear Regression and the Problem of Spuriousness

Tuesday, 10/15 : Read Chapter 9, "Linear Regression and Correlation", and Chapter 10, "Introduction to Multivariate Relationships", Agresti and Finlay. Pick two examples from Tufte, Chapter 3 and read the corresponding sections.

Option for a Research Design Critique: Why are populations of some post-Communist nations more supportive of democracy than others? To what extent is this variation a function of market economic or governing institutional performance? Evans, Geoffrey and Stephen Whitefield. 1995. "The Politics and Economics of Democratic Commitment: Support for Democracy in Transition Societies". *British Journal of Political Science* 25:485-514. Due two weeks by this date.

Option for a Research Design Critique: Why are parliaments more likely to have a greater proportion of female legislators under proportional representation electoral systems? Schmidt, Gregory D. 2009. "The Election of Women in List PR Systems: Testing the Conventional Wisdom." *Electoral Studies* 28(June):190-203. Due two weeks by this date.

Thursday, 10/17 : *Data analysis assignment distributed in class.*

9 Extensions to the Multiple Regression model

Tuesday, 10/22 : Read Chapter 11, "Multiple Regression and Correlation", Agresti and Finlay. And Chapter 4, "Multiple Regression", Tufte.

Thursday, 10/24 : *Data analysis assignment distributed in class.*

10 ANOVA and ANCOVA

Tuesday 10/29 :Read Chapters 12-13, "Analysis of Variance" and "Analysis of Covariance", Agresti and Finlay

Option for a Research Design Critique Do media frames affect public tolerance of hate speech? An experiment was designed to test the effect: Nelson, Thomas E., Rosalee Clawson, and Zoe M. Oxley. 1997. "Media Framing of a Civil Liberties Conflict and Its Effect on Tolerance". *The American Political Science Review* 91:3 567-583. Due within two weeks of this date.

Thursday 10/31 : *Data analysis assignment distributed in class.*

11 Theory—Model Specification and Complications to Regression Analysis

Tuesday, 11/5 : Read Chapter 14, "Model Building with Multiple Regression", Agresti and Finlay.

Article applying linear regression posted on Blackboard Course Documents, depending on class interest and progress.

Option for a Research Design Critique: Simply put, are teachers overpaid? Applying regression analysis, some researchers at the American Enterprise Institute think so. `http://www.aei.org/papers/education/k-12/assessing-the-compensation-of-public-school-teachers/`. One evaluation of the research points you toward a good angle to pursue: `http://www.nytimes.com/roomfordebate/2012/01/02/are-teachers-overpaid/a-better-way-to-slice-the-data.I` within two weeks of this date.

Thursday, 11/7 : *Data analysis assignment distributed in class.*

12 Modeling Variation in Dichotomous Dependent Variables: Logistic Regression

Tuesday, 11/12 : Read Chapter 15, "Logistic Regression", Agresti and Finlay.

Thursday, 11/14 : *Data analysis assignment distributed in class.*

Option for a Research Design Critique: Should the State encourage new parents to install fans (ceiling fans, for example) in infant bedrooms? Explanations for Sudden Infant Death Syndrome (SIDS) have emerged through recent research, pointing toward the absence of air movement as a source of the syndrome. One study applying logistic regression identifies ceiling fans as a factor that dramatically reduces the risk. Coleman-Phox, K., R. Odouli, and L. De-Kun. 2008. "Use of a Fan During Sleep and the Risk of Sudden Infant Death Syndrome" *Archives of Pediatrics and Adolescent Medicine* 162(10):963-968. `http://archpedi.ama-assn.org/cgi/reprint/162/10/963`. Click "PDF" in the upper left side of window to download a copy of the article.

13 Multi-Dimensional Scaling and (K-Means) Cluster Analysis

Tuesday, 11/19 : Read selections from Jacoby, "Data Theory and Dimensional Analysis", posted at *Nota Bene*. And see application on assessing voter coalitions in Brady, "Trust the People" from *The Unfinished Election of 2000*, posted at *Nota Bene*.

Thursday, 11/21 : *Data analysis assignment distributed in class.*

14 Multi-Dimensional Scaling and Cluster Analysis, continued

Tuesday, 11/26 : Read excerpts from Aldenderfer and Blashfield, "Cluster Analysis", posted at *Nota Bene*.

Thursday, 11/28 : Thanksgiving break.

15 Cluster Analysis, continued, class wrap-up

Tuesday, 12/3 : No reading assigned. Course evaluation and wrap-up.

Thursday, 12/5 : Lab analysis work day.

Final Examination Period — Research Paper Presentations

When and Where *You are responsible for verifying the time of the final examination with the Registrar's calendar.* You will present your research paper to the class. Your research paper is due at the beginning of the final examination time slot for the class; it should be posted to Blackboard Course Documents.