

University of Notre Dame  
Political Science 386  
Quantitative Political Analysis  
Spring 2004

Coleman-Morse Computer Lab  
MW 3:00-4:15

Michael Coppedge  
238 Hesburgh Center

coppedge.1@nd.edu  
631-7036

In this course you will learn to understand the most common statistical techniques used in political science and acquire the skills necessary to use these techniques and interpret their results. Mastery of these techniques is essential for understanding research on public opinion and voting behavior, electoral studies, comparative and international political economy, the causes of war, regime change, legislative success, and many other topics. In this class you will read examples of quantitative research on democracy, public opinion, and elections. But more than that, you will actually participate in research by replicating what others have done and by doing original analyses yourself. Therefore, in addition to getting acquainted with some sophisticated research on interesting political topics, you will acquire many of the skills you need to carry out quantitative research on your own. These skills include writing advanced spreadsheet programs; finding, downloading, and "cleaning" datasets; doing the most common sorts of analyses in a statistical package; writing short reports on your findings; and reporting your findings orally and graphically.

The reading load is intentionally light so that you have time to experiment with the abundant data at your disposal, and to do a good job on the many small assignments. Several are ungraded exercises. Nine of them are graded assignments, with the following weights and due dates:

- 10% Feb. 2: Code 4 democracy variables for 20 countries.
- 5% Feb. 9: Calculating means and standard deviations of democracy ratings using Excel.
- 10% Feb. 16: Write your own descriptive report based on survey data, using SPSS.
- 5% Feb. 18: Problem set on standard errors and sample size.
- 5% Mar. 17: Calculate and interpret Chi-square for 2 tables.
- 5% Mar. 22: Replicate Tufte's table 1.1.
- 10% Mar. 24: Calculate inter-coder correlation matrix in Excel using our democracy data.
- 10% Apr. 21: Replicate a multivariate regression using SPSS.
- 15% Apr. 28: Revised multivariate regression analysis of your choice.

There will be a final exam in class, at the computers, which will count as 25% of the grade. I reserve the right to lower your final grade by a sign if you barely participate in class or if your assignments are often late. Assignments must be turned in before the class in which they are due.

Two textbooks are available for purchase at the bookstore in the CAPP section:

- Janet Buttolph Johnson, Richard A. Joslyn, and H.T. Reynolds, *Political Science Research Methods*, 4th edition (Congressional Quarterly Press, 2001) ISBN 1-56802-329-4
- Darren George and Paul Mallery, *SPSS for Windows Step by Step: A Simple Guide and Reference*, 11.0 Update, 4/E (Allyn and Bacon, 2003) ISBN: 0-205-37552-9

Johnson et al. is required; George and Mallery is recommended. Most other readings are available through the library's e-journal locator. Any exceptions are explained below.

My website will have a page of links to the on-line datasets we will be using. My homepage address is <http://www.nd.edu/~mcoppedg/crd>.

Some additional references, not assigned below, that you may find useful include:

- Neil A. Weiss, *Introductory Statistics*, 6th ed. (Addison-Wesley, 2002).
- Wonnacott and Wonnacott, *Introductory Statistics*.
- Michael Lewis-Beck, *Applied Regression: An Introduction*.
- A good web reference work on statistics: <http://www.statsoft.com/textbook/stathome.html>
- A good web reference work on SPSS: <http://sbm-connect.tees.ac.uk/webstat/content%20pages/chapter1/CHAPTER1.HTM>

## Schedule of Readings and Assignments

January 14: Getting to Know You:

Demonstration of a Survey

We will fill out and analyze, in class, a brief, fun survey to get a feel for what is possible.

- No assigned reading.

Exercise for January 22: Come prepared to debate a pre-defined position.

January 19: Politics and Science

Can politics be studied scientifically and quantitatively? A debate.

- Janet Buttolph Johnson, Richard A. Joslyn, and H.T. Reynolds, *Political Science Research Methods*, 4th edition (Congressional Quarterly Press, 2001), chapters 1 & 2. This book is abbreviated as PSRM below.

Exercise for January 22: Come prepared with a hypothesis for refinement during class discussion.

January 21: Political Hypotheses and Concepts

You will present your hypothesis. The class will talk about how it might be quantified and tested.

- PSRM chapter 3
- W. Phillips Shively, *The Craft of Political Research*, 4<sup>th</sup> ed. (Prentice-Hall, 1998),

chapters 2-3. This will be in the Reserve Room and e-reserves.

- Recommended: Charles A. Lave and James G. March, *An Introduction to Models in the Social Sciences* (Harper and Row, 1975), chs 1-3 (1-84).

Part One: Description (Univariate Analysis)

January 26: Measuring Democracy and Other Political Phenomena

You will get acquainted with the basic techniques for measuring political phenomena, using democracy as an example.

- PSRM chapter 4
- Shively chapters 4-5. This will be in the Reserve Room and e-reserves.
- Michael Coppedge and Wolfgang Reinicke, "Measuring Polyarchy," *Studies in Comparative International Development* 25:1 (Spring 1990): 51-72 .
- Recommended: Robert A. Dahl, *Polyarchy* (Yale UP, 1971).

Exercise for January 28: Replicate the coding of the four polyarchy scale variables for five countries using information from 2000. Guidance to sources will be provided in class.

January 28: Using a Spreadsheet to Rate Countries

No additional reading. You will bring in your preliminary coding of 5 countries. We will compare your ratings with others' and discuss and resolve coding discrepancies and if necessary, revise coding criteria. I will show you how to enter your data into an Excel spreadsheet and how to document coding decisions.

Graded Assignment for February 2 (10%):

Submit spreadsheet in advance with coding of an additional 15 countries (to be assigned), for a total of 20.

February 2: How Many Dimensions Does Democracy Have?

You will learn to use functions in Excel, such as creating a sum and sorting by it. This skill will then enable you to try to scale your ratings. How many dimensions are there?

No required reading.

- Recommended: David Nachmias and Chava Nachmias, *Research Methods in the Social Sciences*, second ed. (St. Martin's Press, 1981), chapters 2, 6, and 15;
- Recommended: John P. McIver and Edward G. Carmines, *Unidimensional Scaling* (Beverly Hills: SAGE, 1981).

February 4: The Fuzzy Center: Measures of Central Tendency and Dispersion

You will learn to calculate the mean, mode, median, and standard deviation in Excel, and to generate and interpret histograms. I will demonstrate the Rice Virtual Lab in Statistics (RVLS) simulation on Mean and Median [http://www.ruf.rice.edu/~lane/stat\\_sim/descriptive/index.html](http://www.ruf.rice.edu/~lane/stat_sim/descriptive/index.html). You should experiment with this on your own, too.

- Arend Lijphart, "Party Systems: Two-Party and Multiparty Patterns,"

chapter 5 in his *Patterns of Democracy* (Yale UP, 1999), pp. 62-89. On e-reserves.

- PSRM chapter 11 (305-327 only).

Graded Assignment for February 9 (5%):

Using your spreadsheet, calculate the mean and standard deviation of the component and summary democracy ratings for each country.

February 9: Mining the Internet for Quantitative Political Data

You will learn to convert Excel and other data to SPSS with the SPSS Data Wizard and DBMS-Copy; generate descriptive statistics and graphs in SPSS. I will show you how to access other datasets on the Internet and at the Lab for Social Research, how to use codebooks, and how to recode and transform data.

Exercise: Before class, spend about 45 minutes playing with NationMaster.com

<http://www.nationmaster.com/index.php>

and following the links to data archives found at

<http://www.nsd.uib.no/cessda/namer.html>.

- Recommended: Darren George and Paul Mallery, *SPSS for Windows Step by Step: A Simple Guide and Reference*, 11.0 Update, 4/E (Allyn and Bacon, 2003), chapters 3-4, 6-7. Below this will be referred to as "George and Mallery."

February 11: Reporting Survey Data  
We will discuss a landmark survey of global attitudes as an example of a descriptive survey report. Then you will learn to download survey data for your own analysis.

- Pew Global Attitudes Project, “How Global Publics View: Their Lives, Their Countries, the World, America,” (Pew Research Center for the People and the Press (December 2002) [“PewWorldSurvey.pdf” in the course folder.]
- PSRM chapter 10.

Graded Assignment for February 16 (10%):  
Generate your own report based on at least 5 variables from either the Pew survey (if available) or the General Social Survey. It must employ some measures of central tendency and dispersion and include some informative graphics.

February 16: Sampling: Creating Political Microcosms

Sampling is the nearest thing we have to magic. It gives us the ability to make valid inferences about a whole population by examining only a smart part of it. You will learn how to take a random sample and how to calculate the impact of sample size and selection bias on inferences about a population. We will play with the RVLS Sampling Distribution Simulation at [http://www.ruf.rice.edu/~lane/stat\\_sim/sampling\\_dist/index.html](http://www.ruf.rice.edu/~lane/stat_sim/sampling_dist/index.html). You should experiment with this on your own, too.

- PSRM chapter 7
- Edward Tufte, *Data Analysis for Politics and Policy* (Prentice-Hall, 1974), pp. 40-46 on “Election-Night Forecasting.” This will be in the Reserve Room and e-reserves.
- Henry E. Brady, Michael C. Herron, Walter R. Mebane Jr., Jasjeet Singh

Sekhon, Kenneth W. Shotts, Jonathan Wand, “Law and Data: the Butterfly Ballot Episode (Al Gore and George Bush's Not-So-Excellent Adventure),” *PS: Political Science & Politics* 34:1 (March 2001): 59-69. On-line at <http://www.apsanet.org/PS/march01/brady.cfm>

- Recommended: Jonathan N. Wand, Kenneth W. Shotts, Jasjeet S. Sekhon, Walter R. Mebane Jr., Michael C. Herron, Henry E. Brady, “The Butterfly Did It: the Aberrant Vote for Buchanan in Palm Beach County, Florida,” *American Political Science Review* 95:4 (December 2001): 793-810.

Also, on your own, look at Bill Trochim’s webpages on statistical sampling terms: [http://trochim.human.cornell.edu/kb/sampst\\_at.htm](http://trochim.human.cornell.edu/kb/sampst_at.htm) and probability sampling: <http://trochim.human.cornell.edu/kb/sampprob.htm>

Graded Assignment for February 18 (5%): Do the problem set on standard errors and sample sizes that will be handed out in class.

February 18: How Precisely Do We Know What We Know?

This is a quick introduction to some aspects of probability theory, with an emphasis on the highly useful concepts of the normal distribution and confidence intervals. In class, we will play the Normal Tool Game at the Utah Virtual Lab:

<http://www.psych.utah.edu/stat/bots/game7/Game7.html>

and take a look at the RVLS simulation of confidence intervals:

[http://www.ruf.rice.edu/~lane/stat\\_sim/conf\\_interval/index.html](http://www.ruf.rice.edu/~lane/stat_sim/conf_interval/index.html)

but you should experiment with these on your own, too.

and read:

- PSRM chapters 11 (327-337) and 12 (357-63).
- Michael Coppedge, "The Dynamic Diversity of Latin American Party Systems," *Party Politics* 4:4 (October 1998): 547-568. On e-reserves.
- Recommended: George and Mallery chapter 11.

Part Two: Explanation (Bivariate Analysis)

February 23: Scatterplots

Scatterplots are a beautiful way to get a quick, intuitive idea about the relationship between two continuous variables and the possible reasons for the relationship. You will read an influential example with many scatterplots and learn to produce them in various ways in Excel and SPSS.

- Robert D. Putnam, Robert Leonardi, Raffaella Y. Nanetti, and Franco Pavoncello, "Explaining Institutional Success: The Case of Italian Regional Government," *American Political Science Review* 77:1 (March 1983): 55-74.
- Recommended: George and Mallery chapter 5.

Exercise for February 25: Generate and interpret three scatterplots using any relevant dataset.

February 25: Scatterplot Presentations  
Give an oral presentation presenting and interpreting one scatterplot using any relevant dataset.

March 1: Did All Good Things Go Together in the Twilight of the Soviet Union? Cross-Tabulation  
The real power of surveys comes from cross-tabulation: comparing answers to one question with answers to another. You will read a politically important example of this and learn to do interpret such tables and

produce them yourself.

- Ada W. Finifter and Ellen Mickiewicz, "Redefining the Political System of the USSR: Mass Support for Political Change," *American Political Science Review* 86 (December 1992): 857-874.
- Recommended: George and Mallery chapter 8.
- Recommended: Benjamin I. Page and Robert Y. Shapiro, "The Effects of Public Opinion on Policy," *American Political Science Review* 77:1 (March 1983): 175-190; PSRM12 (388-50 only)

Exercise to present on March 3 (10%): generate and interpret some cross-tabs using General Social Survey data, the 2000 American National Election Study, or the 2002 Global Attitudes Project data (if available).

March 3: Cross-tab Presentations  
You will present your own discoveries and hear about others'.

\* \* \* Spring Break \* \* \*

March 15: Interpreting Tables–Strength of Association vs. Statistical Significance  
Often it is useful to calculate a summary statistic to measure the strength of association in a table. However, some relationships that seem strong may be accidents, the result of chance. How can we find out how much of a danger a chance association is?

- PSRM chapter 12 (pp. 350-357 only).
- re-read Finifter and Mickiewicz to see how the gamma and tau-c coefficients were used.

In class, we will do the RVLS simulation of Chi-square test of deviations at

[http://www.ruf.rice.edu/~lane/stat\\_sim/chisq\\_theor/index.html](http://www.ruf.rice.edu/~lane/stat_sim/chisq_theor/index.html) and the RVLS

Contingency-table simulation at [http://www.ruf.rice.edu/~lane/stat\\_sim/contingency/index.html](http://www.ruf.rice.edu/~lane/stat_sim/contingency/index.html), but as usual, you should experiment with these on your own, too.

- Recommended: George and Mallery chapter 8.

Graded Assignment for March 17 (5%):

Calculate appropriate strength-of-association statistics and do Chi-square significance tests on 2 tables you did before and write a brief (a few sentences) interpretation of each.

March 17: Which Average is Bigger? Really.

There are other techniques for figuring out whether two averages are really different from each other, as opposed to accidentally different. Here you will learn about difference of means tests and the ANalysis Of VAriance (ANOVA). In class, we will try out the Visual ANOVA simulation at the Utah Virtual Lab: <http://www.psych.utah.edu/stat/introstats/anovaflash.html> and do the RVLS simulation on Repeated Measures.

- Edward Tufte, *Data Analysis for Politics and Policy* (Prentice-Hall, 1974), chapter 1 (pp. 1-30). This will be in the Reserve Room and e-reserves.
- PSRM chapters 12 (363-69) and 13 (393-405)
- Recommended: George and Mallery chapters 9 and 12.

Graded Assignment for March 22 (5%):

Replicate Tufte's Table 1-1 (p. 23) on the impact of auto inspections on death rates using the data under "Density" at the bottom of the table. Calculate an appropriate measure of association. You will have to enter the data yourself.

March 22: Which Things Vary Together? Correlation

Correlation coefficients are very common statistics for the association among two variables. You will learn to create them and interpret them correctly.

- Look over Putnam et al., "Explaining Institutional Success," again to absorb the significance of the correlation coefficients.
- Jeffrey A. Segal and Albert D. Cover, "Ideological Values and the Votes of United States Supreme Court Justices" *American Political Science Review* 83:2 (June 1989): 557-565. (e-journal)
- Recommended: George and Mallery chapter 10
- PSRM chapter 12 (377-83).

Graded Assignment for March 24 (10%):

design a spreadsheet correlation program and calculate the inter-coder correlation matrix for any variable in our democracy data. Check your work with SPSS.

March 24: Correlation Presentations  
After trouble-shooting your correlation spreadsheets, we will assess the inter-coder reliability of our polyarchy scale and discuss its interpretation.

March 29: How Much Democracy Can \$1,000 Buy? Simple Regression  
A more useful way to assess the relationship between two variables is to estimate a regression. You will read an influential example from democratization research, which will also give us a reason to discuss non-linear relationships and how to measure and test them. In class, we will do the RVLS simulations of Regression by Eye, Components of r, and the Transformations Demo, but you should experiment with them on your own, too.

- R.W. Jackman, "On the Relation of Economic Development and Democratic Performance," *American Journal of Political Science* 17 (1973): 611-21.
- PSRM chapter 12 (369-377 and 383-85)
- Recommended: George and Mallery chapter 15.

Exercise: Study my Excel spreadsheet on bivariate regression until you understand what it is doing. Use it to do a regression, using any of the datasets we have found, and check your results with SPSS.

March 31: Simple Regression Presentations

After trouble-shooting your bivariate regression spreadsheets, you will present the bivariate relationships you would like to test, from any of the datasets we have at hand.

Part Three: Holding Things Constant (Multivariate Analysis)

April 5: Multivariate regression  
Sometimes what appears to be a relationship between Y and X is really a relationship between Y and Z in disguise. Multivariate regression enables us to hold Z (in fact, often many Zs) constant so we can tell whether the alleged relationship between Y and X is still there.

- PSRM chapter 13 (405-412)
- James L. Gibson, "Truth, Justice, and Reconciliation: Judging the Fairness of Amnesty in South Africa" *American Journal of Political Science* 46:3 (September 2002): 540-556. (e-journal)
- Recommended: George and Mallery chapter 16.

April 7: Interactions in Regression  
We use interactions to model situations in which the impact of X on Y depends on some other condition. Here is an interesting example that asks whether the number of successful parties in a country is the result of social cleavages, election laws, or an interaction between them:

- Octavio Amorim Neto and Gary W. Cox, "Electoral Institutions, Cleavage Structures, and the Number of Parties," *American Journal of Political Science* 41:1 (January 1997): 149-174. The dataset is available for replication at <http://dodgson.ucsd.edu/lij/pubs/>. (It's in SAS format and will need to be converted.)

\* \* \* Easter Break \* \* \*

April 14: Multivariate Regression and the Impact of Campaign Spending  
To make sure you have mastered the concept of multiple regression, we will discuss another example. This one assesses the impact of incumbent's spending and challenger "quality" on congressional election races. I will also show you how to generate and interpret residuals and partial plots and explain why you should do this.

- Donald Philip Green and Jonathan S. Krasno, "Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections," *American Journal of Political Science* 32:4 (November 1988): 884-907. The dataset is available for replication at <http://research.yale.edu/vote/cong78.html> (in ascii format).

April 19: Logistic regression and the Probability of Being a Democracy  
 Ordinary multiple regression is fine when you are trying to explain something that is continuous. But when your dependent variable is an either-or, yes-or-no kind of thing, a different estimator is required: logistic regression. Here you will read a prize-winning example that models the probability that a country is a democracy or a dictatorship. The interpretation of logistic regression coefficients is especially tricky, so I will show you how to create a spreadsheet to translate coefficients into probabilities.

- Adam Przeworski, Michael Alvarez, José Antonio Cheibub, and Fernando Limongi, "What Makes Democracies Endure?" *Journal of Democracy* 7:1 (January 1996): 39-55. Their dataset is available in the course folder as "ACLP.sav"
- PSRM chapter 13 (412-430)
- Recommended: George and Mallery chapter 25

Graded Assignment for April 21 (10%):

Replicate one table from either Amorim Neto and Cox or Green and Krasno using SPSS.

Exercise for April 21 or 26: Do your own multivariate regression analysis (logistic or not) and interpret it. Prepare a presentation of your findings. You will have an opportunity to revise this before it is graded.

April 21: Present and discuss your own multivariate regression analyses.

April 26: Present and discuss your own multivariate regression analyses.

Graded Assignment for April 28 (15%):

Revise your multivariate analysis based on constructive criticism from class and write up your interpretation of it.

April 28: Glimpses of What Lies Beyond  
 A quick tour of where you can go from here: factor analysis, log-linear models, time-series analysis, pooled cross-sectional time-series analysis, ARIMA models, and structural equations. You will not be responsible for this material on the final exam.

I would be happy to hold a review session before the final exam.

May 5, 4:15-6:15 pm: Final exam (25 percent of the grade)  
 There will be no final paper.