This course is the second in a two-course sequence aimed at helping Evans School MPA students become informed users and critical consumers of research and statistical analyses. This course introduces the application of probability, hypothesis testing, and confidence intervals to multivariate models in the context of policy and management research. By the end of this course, you will:

- Formulate answerable research questions that address complex policy questions.
- Be aware of the conditions necessary to establish causal relationships on a given outcome, emphasizing the need to disentangle the effects of multiple factors.
- Recognize the implications of research design choices, randomization, concept measurement, and good data collection for the validity and reliability of research results.
- Discriminate among data collection methods appropriate to answer a given research question, such as surveys, focus groups, key informant interviews, administrative data, or other methods.
- Understand the mechanics, assumptions, and interpretation of regression models to policy or management questions, how to use regression models for both prediction and hypothesis testing, and the assumptions behind and possible "fixes" for problems with models.
- Read and analyze empirical studies
- Produce a useful multivariate empirical analysis for a non-statistician, including clear data presentation and the graphical display of data.
**Texts:**
We have two texts for this course, one required and one optional. If you took PBAF 527 last quarter, you should already have Statistics 9th Edition by James T. McClave and Terry Sincich (Prentice Hall: Upper Saddle River, NJ, 2003). In the syllabus, I refer to it as M&S. The optional text is The Basics of Social Research, Second Edition by Earl Babbie (Babbie in the syllabus).

**Software:**
You will also need SPSS or Excel to complete some of the class assignments. Recommended resources for SPSS and Excel are Adventures in Social Research: Data Analysis Using SPSS 11.0/11.5 for Windows and Excel Manual: Statistics. Another recommended resource for SPSS is Doing Data Analysis With SPSS 10.0 by Robert H. Carver and Jane Gradwohl Nash. The texts are available at the University Bookstore and you can get SPSS and Excel from the University Bookstore Computer Outlet. You can use the SPSS and Excel software that are available in the computer lab at the Evans school or at the Center for Social Science Computation and Research (CSSCR) (in the basement of Savery Hall). I encourage you to use the SPSS software.

**Teaching Notes:**
My teaching notes for the class will be available on the class website. It will be beneficial if you access/print them prior to the lecture. They will be available for downloading on the Tuesday before the class session.

**Some tips:**
- Do the reading before you come to class. Hearing the material after having read it will help things make sense to you. I plan class exercises and lecture with the assumption that you’ve done the reading.
- Attend class regularly, keep up with your assignments, and frequent the weekly review sessions. If you feel like you do not understand some concept, ask Gretchen or myself for help!
- A good way of learning statistics is to discuss the material with someone else. To that end, I recommend you form a study group with whom you work on problem sets. Set a time regular time to meet. Ask questions among yourselves. Seeing material from the perspectives of others will help you better to formulate your understanding of the material. You should hand in your own work after having reviewed your responses to the problems with your group.
**Course requirements:**
The course requirements include five homework sets, two in-class exams (open books and notes) and a final project. The purpose of the two exams is to help diagnose your progress in learning the mechanics and interpretation of regression. The policy report allows you to consolidate your learning about regression models, apply what you are learning to a policy context, and learn to communicate your results to a nontechnical audience.

About the homework: homework sets are graded on a 0-10 point scale and are meant to facilitate your learning the material. The homework (with the exception of Assignment 1) will be available at the course Web site on the day noted in the course outline. Answer sheets will be handed out on the due date of each assignment, so no late homework will be accepted. If you cannot hand in the homework on time, I recommend that you complete it anyway so that you learn the material.

**Grading:**
Grades will be based on the five homework sets, two exams and project as shown below:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework assignments</td>
<td>10%</td>
</tr>
<tr>
<td>In-Class Exam I (April 27-open book and notes)</td>
<td>25%</td>
</tr>
<tr>
<td>In-Class Exam II (May 25-open book and notes)</td>
<td>35%</td>
</tr>
<tr>
<td>Policy Report</td>
<td>30%</td>
</tr>
<tr>
<td>Proposal due May 4 in class</td>
<td></td>
</tr>
<tr>
<td>Final project due June 8 in class</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Topic</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>1. Mar 30</td>
<td><strong>Correlation, Causality &amp; Linear Relationships</strong>&lt;br&gt;Purposes of Research; Bivariate Associations; Causation&lt;br&gt;<strong>Regression Analysis &amp; the Research Process</strong>&lt;br&gt;Types of Models&lt;br&gt;Simple linear regression</td>
</tr>
<tr>
<td>3. Apr 13</td>
<td><strong>What about other causal factors?</strong>&lt;br&gt;Multiple Regression Model&lt;br&gt;R², adjusted R²&lt;br&gt;Hypothesis Testing; Confidence Intervals for Parameter Estimates</td>
</tr>
<tr>
<td>4. Apr 20</td>
<td><strong>Interpreting Results</strong>&lt;br&gt;P value, Errors, Goodness of Fit, F Test, Estimation, Prediction&lt;br&gt;Omitted variables; Classical Assumptions&lt;br&gt;The Hunt for Data&lt;br&gt;Exam Preview</td>
</tr>
<tr>
<td>5 Apr 27</td>
<td><strong>How do we choose our model?</strong>&lt;br&gt;The role of theory&lt;br&gt;Dummy variables&lt;br&gt;Interaction Terms&lt;br&gt;<strong>EXAM I</strong></td>
</tr>
<tr>
<td>6. May 4</td>
<td><strong>How do we choose our model?</strong>&lt;br&gt;Examples of Dummy variables and Interaction Terms</td>
</tr>
<tr>
<td>7. May 11</td>
<td><strong>What are some problems with our model?</strong>&lt;br&gt;<strong>The Regression User’s Guide</strong>&lt;br&gt;Detecting and fixing regression Problems</td>
</tr>
<tr>
<td>8. May 18</td>
<td><strong>Policy Report Workshop</strong>&lt;br&gt;So, you want to collect data&lt;br&gt;Choosing the right method for the question.&lt;br&gt;Pitfalls to Avoid</td>
</tr>
<tr>
<td>9. May 25</td>
<td><strong>Policy Report Question and Answer Session</strong>&lt;br&gt;<strong>Evaluating Empirical Studies</strong>&lt;br&gt;<strong>EXAM II</strong></td>
</tr>
<tr>
<td>10. June 1</td>
<td><strong>Discussion of Projects, Course Evaluations</strong></td>
</tr>
</tbody>
</table>

**Final Project Due**  Wednesday June 8, 2005 by 12 noon
In this project, you will structure an empirical question, search for the data that helps you to answer that question, strategize about modeling approaches, carry out an analysis, and interpret the results for a client. You will present your findings in written report, a "memo" to a specific client who is known to be a non-statistician. You can continue to work on topics from previous quarters, or you can choose a new subject area. In choosing a topic, keep in mind the following questions:

- Is this a topic that can be analyzed with statistical tools?
- Is there adequate information available for the analysis?
- Is the topic interesting enough to hold my interest for a quarter?
- Will I expand my skills and knowledge by completing this project?

You will need to find a partner with whom to do the project. You will receive a team grade--be sure that your expectations and goals are compatible.

Data
Get your data from government documents, publications, inside contacts, your own survey, CSSCR, or wherever you can find it. For this project, you should have at least 40 cases (observations) and 8 or 10 variables. You will need at least 1 outcome variable and 4 or 5 explanatory variables in your final model. If your data are not available electronically, then you may want to enter them into a spreadsheet program such as Excel before analyzing it in Excel or SPSS. Gathering and preparing data takes lots of time. To avoid a crunch, start early in the quarter. This is your opportunity to find a data set that is relevant to your interests. You may use the 2004 WAPOP in conjunction with other data but I discourage you from using it exclusively.

Policy Project Concept
Your project concept is due on April 13. The project concept will state who you are working with and what types of questions you are seeking to answer. It need not be more paragraph.

Proposal
The proposal for your final project is due on May 4. Make your proposal a memo to a real or fictitious client. Briefly outline the topic, your research hypotheses, your data sources, and the outcome and explanatory factors. It should also contain tables, charts, and graphs that illustrate the nature of the variables. For each variable you should be sure to note how many observations you have, and how they are measured (are they categorical or quantitative/continuous?) You should note how many observations are available for regression analysis. A typical project proposal will consist of 2 pages of text and 2 pages of tables/charts/graphs.
Final project
The final report, due June 8, should be a memo of 8-10 single-spaced pages in length (including graphics and tables) written to your non-statistician client. The report should demonstrate your understanding of the course material in the context of your chosen example and in language understandable to a lay-person. Use footnotes and appendices for the technical information that someone with a more technical bent would need to evaluate what you have done. Remember to use descriptive statistics and graphics to complement your analysis. These may be very valuable to your client and other audiences. The memo will be graded on content (which includes the quality of your research question, creativity in creating your data set, and how well your data relates to your question), application of statistical tools, clarity of writing and graphical presentation. You are encouraged to make full use of statistical software, word processing, and graphics packages to put your results in a readable form.

Tips
• Talk to others about your projects and data challenges.
• Describe the general issue and topic in the memo before giving technical and statistical information. In some cases this would be a literature review, in others a discussion of the problem in general. This section presents the motivation for the rest of the report.
• Provide an executive summary of one-page or less at the beginning of your policy report. This is not your introduction.
• Try alternative models and, if appropriate, include some of them in your report.
• Use descriptive statistics and graphics for visual presentation of information.
• Use words and phrases; never use meaningless variable labels in your memo.
• Discuss issues relevant to your topic -- not general problems with regression analysis.
• Check spelling and grammar.
• Appearance counts -- work on layout.
• Translate your results into something understandable to non-technical person.
• Use footnotes and appendices for the technical information that someone with a more technical bent (e.g. your instructor) would need to carefully evaluate your analysis.