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.
- Understand the mechanics, assumptions, and interpretation of regression models, how to use regression models for both prediction and hypothesis testing, and the assumptions behind the models.
- Read and analyze empirical studies.
- Produce a useful multivariate empirical analysis for a non-statistician.

**Texts:** The primary text for the course is *Statistics (9th Ed.)* by James T. McClave and Terry Sincich (Upper Saddle River, NJ: Prentice Hall, 2003). I will also make available a few readings during the quarter on research methods and on regression analysis for dichotomous variables.

**Software:** You will need to use SPSS to complete some of the class assignments and the final project. A recommended resource for SPSS is *Doing Data Analysis With SPSS 10.0* by Robert H. Carver and Jane Gradwohl Nash. You can get SPSS from the University Bookstore Computer Outlet. You can use the SPSS software that is 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).
Course requirements: The course requirements include five homework sets, two in-class exams (open books and notes), and a final project. The homeworks are graded with a check, check minus, or check plus and are meant to facilitate your learning the material. The homework will be available on the day the previous assignment is due. As answer sheets will be made available on the due date of each assignment, no late homework will be accepted. Grades will be based on the homeworks, exams, and project as follows:

- **Homework assignments** 10%
- **In-Class Exam I** (April 26-open book and notes) 25%
- **In-Class Exam II** (May 26-open book and notes) 35%
- **Policy Project** 30%
  - Project Concept due April 14 in class
  - Formal Project Proposal due May 5 in class
  - Final Project due at 12pm on June 8

### Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Topics</th>
<th>Reading</th>
<th>Due Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Week 1</strong> 3/29, 3/31</td>
<td>Dependence, Correlation, Linear Relationships, Simple Linear Regression</td>
<td>Ch. 2.9, 11.6, begin Ch. 11.</td>
<td></td>
</tr>
<tr>
<td><strong>Week 2</strong> 4/5, 4/7</td>
<td>Simple Linear Regression: Model Assumptions, ( R^2 ), Prediction, Hypothesis Testing</td>
<td>Ch 11</td>
<td>HW#1, 4/7</td>
</tr>
<tr>
<td><strong>Week 3</strong> 4/12, 4/14</td>
<td>Begin Multiple Regression</td>
<td>Begin Ch. 12</td>
<td>Project concept, 4/14</td>
</tr>
<tr>
<td><strong>Week 4</strong> 4/19, 4/21</td>
<td>Continue Multiple Regression: Interpreting Coefficients, Model Assumptions, Inferences About Parameters, Adjusted ( R^2 ), Prediction</td>
<td>Ch.12</td>
<td>HW#2, 4/21</td>
</tr>
<tr>
<td><strong>Week 5</strong> 4/26, 4/28</td>
<td>Continue Multiple Regression: F-Tests</td>
<td>Ch. 12</td>
<td>Exam #1, 4/26</td>
</tr>
<tr>
<td><strong>Week 6</strong> 5/3, 5/5</td>
<td>Continue Multiple Regression: Interaction terms, Squared terms, and other non-linear terms. More on F-Tests</td>
<td>Ch. 12</td>
<td>Formal Project Proposal, 5/5</td>
</tr>
<tr>
<td><strong>Week 7</strong> 5/10, 5/12</td>
<td>Detecting Regression Problems Logistic Regression: Methods</td>
<td>Reading TBA</td>
<td>HW #3, 5/10</td>
</tr>
<tr>
<td><strong>Week 8</strong> 5/17, 5/19</td>
<td>Logistic Regression: Hypothesis Testing. Empirical Studies</td>
<td>Reading TBA</td>
<td>HW #4, 5/17</td>
</tr>
<tr>
<td><strong>Week 9</strong> 5/24, 5/26</td>
<td>Review and Empirical Studies</td>
<td>Articles TBA</td>
<td>Exam #2, 5/26</td>
</tr>
<tr>
<td><strong>Week 10</strong> 5/31, 6/2</td>
<td>Empirical Studies and Further Topics in Quantitative Analysis</td>
<td>Articles TBA</td>
<td>HW #5, 6/2</td>
</tr>
</tbody>
</table>

**FINAL PROJECT DUE BY 12 NOON ON WEDNESDAY JUNE 8.**
Policy Project
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. 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 must have a partner for this project.

Data
Get your data from government documents, publications, inside contacts, CSSCR, possibly your own survey, or wherever you can find it. For this project, you should have at least 40 observations (cases) and 4 or 5 (or more) variables. Hopefully you will have many more than 40 observations. If your dependent variable is dichotomous (a 0/1 variable, e.g. earned a high school degree versus did not earn a high school degree) then you will need 100 or more observations. 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! While we will use the 2004 WAPOP data for classroom purposes and for homework from time to time, I discourage the use of this data for your class project.

Policy Project Concept
Your project concept is due on April 14. 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.

Formal Project Proposal
The formal project proposal, due May 5, will consist of a memo to a real or fictitious client. Briefly outline the topic, your hypotheses, your data sources, and the dependent and independent variables. 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 2-3 pages of text and 2-3 pages of tables/charts/graphs.
Final Project
The final report, due June 8 at 12 pm, 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. 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.
• 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 a 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.