Statistics are not magical. Nor are they always true – or always false. Nor need they be incomprehensible. Adopting a critical approach offers an effective way of responding to the numbers we are sure to encounter. Being critical requires more thought, but failing to adopt a critical mind-set makes us powerless to evaluate what others tell us. – Joel Best, “Telling the Truth about Damned Lies and Statistics,” The Chronical of Higher Education, May 4, 2001.

Overview. This course is designed to prepare students for advanced quantitative methodology courses required of doctoral students. The course begins by reviewing descriptive statistics and data presentation techniques. In preparation for the study of inferential statistics, the next section of the course covers the basics of probability. A solid grounding in probability is necessary to understand how and why statistical techniques work. Building on that foundation, the heart of the course is a rigorous introduction to statistical inference: sampling theory, confidence intervals, and hypothesis tests. The final section of the course is an introduction to regression analysis, with an emphasis on interpretation of regression results, using examples from recent research. This course is part of a two semester sequence; the second semester is Quantitative Methods II, which is a more advanced and detailed treatment of regression analysis and related topics.

Learning Objectives. The main objective of the course is that students will develop a conceptual and practical understanding of how to learn from social science data; specifically:

- Students will learn how to calculate and interpret descriptive statistics;
- Students will have a grounding in probability and understand how to work with random variables, univariate probability distributions, and correlation between variables;
- Students will understand the theory behind inferential statistics that supports drawing conclusions about a population from a sample;
- Students will be able to calculate confidence intervals and conduct basic hypothesis tests for means, proportions, and variances;
- Students will understand Type I and Type II errors and will be able to calculate the power of the test;
- Students will be able to read and interpret basic Ordinary Least Squares regressions as found in many academic journal articles.
Prerequisites. The topics discussed in this course are inherently mathematical. Students taking this course should be very comfortable with college algebra. Some calculus will be used, but no knowledge of calculus is assumed or required.

Computing. This is a course about concepts, not software. Nevertheless, we will use software at times to reduce the computation burden. *Stata*, version 13.1, is the official software for the class, but any recent version of Stata will work nearly as well. *You do not need to buy Stata.* It is available on the Rutgers on-line system (apps.rutgers.edu) and in the Robeson Computer Lab. However, if you wish to use Stata on your own equipment, a special discounted version will be available to students through the “grad plan.” You can get a six-month license for “Small Stata” (the student version) for only $35. This version is limited to 1,200 observations, but it will be adequate for this course. However, if you are going on to Quantitative Methods II in the Spring and/or have an interest in doing empirical work, it would make sense to order at least Stata/IC (one year license, $98; perpetual license, $189). To order, please contact StataCorp directly: Phone: 800-782-8272 (Monday through Friday 8:00 to 5:00 Central Time). Be sure to ask for the “grad plan.” If ordering online, use your Rutgers email address to verify affiliation with the university. More information is available via a link from the course web site in Sakai.

Miscellaneous. A calculator is a necessity. It does not have to be fancy and it does need graphics or programming capabilities, but it is useful to have the following functions: \( \ln(x) \), \( e^x \), \( x! \), and \( y^x \) (logarithms, exponents, factorials, and powers). Usually any calculator described as “scientific” will have these functions. Several smart phone apps are available. Good choices are available for as little as $10. For this class, graphing capability and programmability will not be needed.

Grading and Requirements:

Problem Sets: The only sure way to learn the material presented in this course is to work on problems that reinforce the readings and lectures. Thus, there will be a short problem set due almost every week (see the schedule below for exact due dates). Except in unusual circumstances, *late problem sets will not be accepted* because the correct timing of the work is important in the learning process and because the answers will be discussed in class. However, *the lowest two problem set grades will be dropped*, allowing the student some flexibility and margin for error.

Tests: There will be two midterm tests during the semester on Oct. 2 and Nov. 6. The first covers fundamentals of data, descriptive statistics, correlation, and probability. The second covers sampling theory and hypothesis testing. The final exam, scheduled for Dec. 18, is comprehensive. The tests are *open-book, open-note*. A calculator is a necessity, hopefully one with which you are familiar. *Laptop computers are not permitted during the test.* Mark your calendar now, because in fairness to other students, I cannot create make-up tests or reschedule tests for any one person.
Grades will be assigned as follows: Letter grades will be determined based on the overall course average, rounded to the nearest whole number. Only certain letter grades are available for graduate course work at Rutgers-Camden. The translation of a numeric grade to a letter grade will be done as follows:

A 90-100
B+ 85-89
B  80-84
C+ 75-79
C  70-74
F  0-69

In determining the course average, assignments will be weighted as follows:

- Problem Sets (lowest two dropped): 40 percent
- 2 Midterms @ 15 percent each: 30 percent
- Final exam: 30 percent

At any point in the semester, you can see your current course grade in Sakai based on the graded work up to that point. You can also figure out what grade you need on the final exam to get any specific final grade.

Incompletes: Generally speaking, the material in this course is best learned as a single unit. I will grant incompletes only in cases where a substantial change in life circumstances occurs that is beyond the control of the student, and then only with appropriate documentation.

**Ground Rules for the Course.**

**Attendance.** Attendance is entirely optional. Having said that, you are responsible for everything covered in class whether it was covered in the readings or not. You are also responsible for any announcements made in class. For most people, attendance is a necessary condition for learning the material. The PowerPoint slides are not a substitute for attending the class, because the slides are not self-contained – they are props to give structure to my lectures and our class discussions. If you need to miss a class, be sure to get notes and a recap from a classmate.

**Study Groups.** You are encouraged to form and participate in a regular study group. Many students over the years have found the study groups to be very helpful. Study groups are permitted and encouraged to work on the problem sets together. However, each individual student should write up his or her own answer to hand in, based on his or her own understanding of the material. Do not hand in a copy of another person’s problem set, even a member of your own group. Writing up your own answer helps you to internalize the group discussions and is a crucial step in the learning process.
Academic Integrity. Violations of academic integrity include cheating on tests or handing in assignments that do not reflect your own work and/or the work of a study group in which you actively participated. I have a policy of zero tolerance for cheating. Violations will be referred to the appropriate university authorities.

Schedule and Readings. These are the planned readings, test dates, and assignment due dates. Always check the online version of the course schedule in Sakai for updates to the schedule. Also, be sure to note the date of the quizzes and final exam.

<table>
<thead>
<tr>
<th>Date</th>
<th>Title of Class</th>
<th>Topics</th>
<th>Readings</th>
<th>Assignments Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sep. 4</td>
<td>Descriptive Statistics</td>
<td>• Types of Data</td>
<td>• Jargowsky and Yang, <em>Descriptive and Inferential Statistics</em>, Section I, pp 1-16.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Variability</td>
<td>• Gene Zelazny, <em>Say it with Charts</em>, Introduction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Skewness and Kurtosis</td>
<td>• Open Intro Text</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Correlation and Relationships</td>
<td>o Section 1.2, “Data Basics”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Graphic Techniques</td>
<td>o Section 1.6, “Examining Numerical Data”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o Section 1.7.1, “Contingency Tables and Bar Plots”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>o Section 1.7.2, “Row and Column Proportions”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Introduction to Stata Interface</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• <a href="https://www.youtube.com/watch?v=L8ilj_8lhRc">https://www.youtube.com/watch?v=L8ilj_8lhRc</a></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td><strong>Effective Presentation of Data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topics</td>
<td>- Identifying the informative contrast</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Effective communication with tables and graphs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readings</td>
<td>- Gene Zelazny, <em>Say it with Charts</em>, Section 1: Choosing Charts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Tufte, <em>The Decision to Launch the Space Shuttle Challenger</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Watch the Challenger disaster <a href="https://www.youtube.com/watch?v=AfnvFnzs91s">https://www.youtube.com/watch?v=AfnvFnzs91s</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Tufte, “The Cognitive Style of Powerpoint”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assignments Due</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- <strong>Problem Set #1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th><strong>Basics of Probability</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics</td>
<td>- Simple and Compound Events</td>
</tr>
<tr>
<td></td>
<td>- Bayes’ Theorem</td>
</tr>
<tr>
<td></td>
<td>- Random Variables</td>
</tr>
<tr>
<td></td>
<td>- Expected Value and Variance</td>
</tr>
<tr>
<td>Readings</td>
<td>- Open Intro Text, Chapter 2, “Probability”</td>
</tr>
<tr>
<td>Assignments Due</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- <strong>Problem Set #2</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th><strong>Probability Distributions</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics</td>
<td>- Discrete probability distributions</td>
</tr>
<tr>
<td></td>
<td>- Continuous probability functions</td>
</tr>
<tr>
<td>Readings</td>
<td>- Open Intro Text:</td>
</tr>
<tr>
<td></td>
<td>- Section 3.3, “Geometric Distribution”</td>
</tr>
<tr>
<td></td>
<td>- Section 3.5.2, “Poisson Distribution”</td>
</tr>
<tr>
<td></td>
<td>- Chapter 3.1, “Normal Distribution”</td>
</tr>
<tr>
<td>Assignments Due</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- <strong>Problem Set #3</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Date</th>
<th>***** Test 1 *****</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Open book, open note.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Bring a calculator!</strong></td>
</tr>
</tbody>
</table>
| Oct. 9 | **Sampling Theory**  
|        | - Basic principles of sampling  
|        | - Sampling error vs. bias  
|        | - The Central Limit Theorem  
|        | **Readings**  
|        | - Open Intro Text:  
|        |   - Section 1.3.1, “Populations and samples”  
|        |   - Section 1.3.3, “Sampling from a population.”  
|        |   - Section 1.4.2, “Three sampling methods”  
|        |   - Introduction to Chapter 4  
|        |   - Section 4.1, “Variability in estimates”  
| Oct. 16 | **Point and Interval Estimates, Hypothesis Testing Framework**  
|        | **Topics**  
|        | - Confidence intervals for means and proportions  
|        | - Confidence intervals for two-sample differences  
|        | - Introduction to hypothesis testing  
|        | **Readings**  
|        | - Open Intro Text:  
|        |   - Section 4.2, “Confidence Intervals”  
|        |   - Section 4.3, “Hypothesis Tests”  
|        |   - Section 4.4, “Examining the Central Limit Theorem”  
| **Assignments Due** | **Problem Set #4**  
| Oct. 23 | **Hypothesis Testing I: Large Sample Tests**  
|        | **Topics**  
|        | - Hypothesis Testing  
|        | **Readings**  
|        | - Open Intro Text:  
|        |   - Section 5.1, “Paired data”  
|        |   - Section 5.2, “Difference of two means”  
|        |   - Section 6.1, “Inference for single proportion”  
|        |   - Section 6.2, “Difference of two proportions”  
|        |   - Section 6.6, “Hypothesis testing for two proportions”  
| **Assignments Due** | **Problem Set #5**
### Oct. 30
**Hypothesis Testing II: Small Sample Tests and Statistical Power**

**Topics**
- Student’s t distribution
- Pooled variance estimator
- The power of the test
- Significance vs. Importance

**Readings**
- Open Intro Text:
  - Section 5.3, “One-sample means with the t distribution”
  - Section 5.4, “The t distribution for the difference in two means”
  - Section 4.6, “Sample size and power”

**Assignments Due**
- Problem Set #6

### Nov. 6

**Test II**
*Open book, open note.*
*Bring a calculator!*

### Nov. 13
**Hypothesis Testing III: More Tests and Tricks**

**Topics**
- ANOVA
- Chi-Squared test
- Bootstrapping

**Readings**
- Open Intro Text:
  - Section 5.5, “Comparing many means with ANOVA”
  - Section 6.3, “Testing for goodness of fit using chi-square”
  - Section 6.4, “Testing for independence in two-way tables”

### Nov. 20
**Hypothesis Testing IV: When the Sampling Distribution is Unknown**

**Topics**
- Bootstrapping
- Non-parametric tests

**Readings**
- Karl Wuensh, “Nonparametric Statistics.”
- Open Intro Text:
  - Section 6.5, “Small sample hypothesis testing for a proportion”

**Assignments Due**
- Problem Set #7
<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Topics</th>
<th>Readings</th>
<th>Assignments Due</th>
</tr>
</thead>
</table>
| Nov. 25*  | **Bivariate Regression**  | - The population regression function  
- The sample regression function  
- Obtaining the estimates  
- Testing the hypothesis that X affects Y | - Tuft, Chap. 3, pp. 65-91   
- *Tuft, Chap 3, pp. 91-end. | - Problem Set #8 |
| Dec. 4    | **Multiple Regression**   | - Isolating the Effect of X on Y  
- Left Out Variable Bias  
- The true meaning of multiple regression  
- Hypothesis testing  
- Goodness of Fit | - Tuft, Chapter 4.  
| Dec. 18   | **** Final Exam *****     |                                                                 |                                                                                              | **** Final Exam *****   |

*Note: Thursday class meeting on Tuesday!*