Course Description

This course offers an introduction to statistics. The course will cover descriptive and inferential statistics including frequency, probability, hypothesis testing, ANOVA, and regression analysis. Special attention throughout the course will be paid to the critical consumption of statistics in the media and the responsible deployment of statistics in research. The course emphasizes comprehension, application, and communication of research findings.

Required materials

- Making Sense of Data Through Statistics 2nd Ed, Nevo
  
  A pdf version of this text is available online at www.ldpress.com for $19.95; printed versions of the text can be ordered for $90

- Additional readings, assignments, lecture slides, and grades will be made available on Blackboard

Course Structure

Overview of expectations, assignments, and exams

We’ll cover a chapter of Nevo (Second Edition) in addition to a variety of assigned readings (posted on Blackboard) each week. Your classmates and I expect you to come to class prepared, meaning that you will have read the week’s assigned readings in advance of our class meeting on Monday each week.

We’ll have two exams and series of problem sets over the course of the semester. In addition, you will work towards a data-based statistical research paper throughout the semester on a topic that is of interest to you. Finally, pop quizzes may appear without notice; these will be used to assist me in assessing your preparation for and comprehension of course materials.

Consultation and collaboration with your peers is encouraged for the completion of problem sets (with the cautionary note that free riders will pay a serious penalty on quizzes, exams, and the final research paper); however, the final research paper must be uniquely yours, with proper attribution of all sources.
Research paper

You will work towards a research paper in this course by meeting several milestones over the course of the semester. The milestone dates and the extent to which these milestones contribute to your grade on this assignment are listed below; the dates are also included in the schedule below.

- Research question and data set (10%), due February 1
- Brief literature review (10%), due February 22
- Summary statistics and methods (10%), due March 22
- Analysis and discussion of findings (10%), due April 5
- Visit to the writing center (5%), due between April 5 and April 19
- Final paper (55%), due April 26

Optional training sessions

The course TA will hold optional training sessions outside of regular class hours to assist those who are interested in becoming more proficient in Excel, SAS, R, and the Microsoft Equation Editor. More (or fewer) sessions may be offered depending on student interest and availability. The times and dates for these optional training sessions will be determined on the second day of class.

Discussion board

We’ll use a discussion board approach to troubleshooting your problem sets. You’ll find the discussion board on Blackboard. I will create a forum for each problem set once it is distributed; feel free to add threads for your specific questions. Like a "leave a penny, take a penny" jar, the success of this forum will depend on your contribution. If everyone is only posing questions and no one is answering questions, then the forum will fail. If you go to the forum to ask a question, take time to look at your classmates’ questions and see if you can help.

Your participation grade will be based on how well you assist your peers’ learning in class and/or in this forum.

The discussion board approach to troubleshooting problem sets mirrors how many statisticians, econometricians, data scientists, and programmers troubleshoot problems IRL. See for example, StackExchange. Note that such forums have codes of conduct based on kindness and clarity of communication. Visit the StackExchange Code of Conduct for guidance; failure to follow this code of conduct will result in a zero for your participation grade.

Grading

Your course grade will be a weighted average of the following aspects of the course:

- Participation (in class and in problem set discussion boards) 5%
- Problem sets and pop quizzes 25%
- Midterm 20%
- Final 20%
- Research paper 30%

Letter grades will be assigned according to your weighted average for the course. Grades will be available in Blackboard as they are recorded.
<table>
<thead>
<tr>
<th>Average Grade</th>
<th>Average Grade</th>
<th>Average Grade</th>
<th>Average Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>87-89% B+</td>
<td>77-79% C+</td>
<td>67-69% D+</td>
<td>93-100% A</td>
</tr>
<tr>
<td>93-100% A</td>
<td>83-86% B</td>
<td>73-76% C</td>
<td>80-82% B-</td>
</tr>
<tr>
<td>90-92% A-</td>
<td>70-72% C-</td>
<td>0 - 62% F</td>
<td>63-66% D</td>
</tr>
</tbody>
</table>

**Late assignments**

I will accept late problem sets without penalty so long as a new due date was arranged in communication with me 48 hours in advance of the original due date. Problem sets submitted after the due date without prior communication will receive a grade of zero.

Because the research paper milestones build on one another, and because I will have a tight turn around time during which to provide you with feedback before the next milestone, late submissions of these assignments will receive a zero.

Emergencies and other extenuating circumstances affecting your ability to turn in your work on time will be considered on a case-by-case basis. Overall, communication with me in advance of the due date will be rewarded while lack of communication is likely to result in a zero.

**Attendance**

Your regular attendance in class is expected. If you anticipate a conflict with an exam or an assignment due date, you must reach out to me as soon as possible but no later than 48 hours prior to the exam or due date to make alternative arrangements. Should you be absent for an exam or an assignment due date without having made prior arrangements, you will receive a zero for the exam or assignment.

As with late assignments, emergencies and other extenuating circumstances affecting your ability to attend class will be considered on a case-by-case basis. Overall, communication with me in advance will be rewarded while lack of communication is likely to result in a zero.

**Academic integrity**

Students are required to comply with the Code of Student Conduct; this includes proper attribution of all sourced materials.

**General Schedule**

Please note that this schedule is tentative and subject to change as the semester progresses. Exam dates and important due dates for the research paper are noted in **bold**. Problem sets will be given on, at most, a weekly basis.

**Week 01, 01/14 - 01/18:** Introduction to data, statistics, & data visualization

- Read Nevo Chapter 1
- Read Nevo Chapter 2

**Week 02, 01/21 - 01/25:** Measures of centrality & variation

- Read Nevo Chapter 3
- Read Strogatz "Friends you can count on"
Week 03, 01/28 - 02/01: Probability
- *Read* Nevo Chapter 4
- *Read* Strogatz "Chances are"
- Due Friday, February 1 Research question and data set

Week 04, 02/04 - 02/08: Discrete probability distributions
- *Read* Nevo Chapter 5

Week 05, 02/11 - 02/15: Continuous probability distributions
- *Read* Nevo Chapter 6
- *Read* Strogatz "The new normal"

Week 06, 02/18 - 02/22: Introduction to hypothesis testing
- *Read* Nevo Chapter 7
- *Read* Caldwell "Four fundamental concepts"
- Due Friday, February 22 Literature review

Week 07, 02/25 - 03/01: Hypothesis testing II & midterm
- *Read* Nevo Chapter 8
- Midterm Friday, March 1st

Week 08, 03/04 - 03/08: Hypothesis testing III
- *Read* Nevo Chapter 9

Week 09, 03/11 - 03/15: Spring Break
- *Note* No class

Week 10, 03/18 - 03/22: Hypothesis testing IV
- *Read* Nevo Chapter 10
- Due Friday, March 22 Summary statistics & plan for analysis

Week 11, 03/25 - 03/29: Chi square tests & ANOVA
- *Read* Nevo Chapters 11 12

Week 12, 04/01 - 04/05: Simple regression
- *Read* Anderson, Sweeney, & Williams Chapter 12
- Due Friday, April 5 Analysis & discussion of findings
- Due Friday, April 5 Make an appointment to visit the Writing Center
Week 13, 04/08 - 04/12: Multiple regression
  • Read Anderson, Sweeney, & Williams Chapter 13

Week 14, 04/15 - 04/19: Multiple regression & econometrics
  • Read Wooldridge "The nature of econometrics and economic data"
  • Read Angrist & Pischke "The experimental ideal"

Week 15, 04/22 - 04/26: Predictive analytics & review
  • Read Kleinberg et al 2015 AER
  • Due Friday, April 26th Final research paper

Final Exam Saturday, May 4, 9-11:15am