

PADP 7120: Data Applications

University of Georgia
Fall 2019
Tu, 6:15-9:00 PM, Gwinnett 121
Professor Anastasopoulos

Office: 3rd Floor Gwinnett Building
Email: ljanastas@uga.edu
Office Hours: TBA

Course Overview and Objectives

This course is an introduction to statistical reasoning and thinking applied to understanding the policy world. The tools that we will learn in this course will help you not foster a better understanding of policy problems more broadly, but will help clarify the way that you think and reason through problems where data is involved. The topics that we will cover include: research design, measurement, causal inference and understanding how we measure and interpret relationships between events and concepts.

Required Text and Readings

The required text for the course:

Textbook: *We will use an open source statistics book—you can download a free pdf copy here:* https://www.openintro.org/stat/index.php?stat_book=os.

Diez, David M., Christopher D. Barr, and Mine Çetinkaya-Ründel. 2016. *OpenIntro Statistics*. 3rd Eds. www.openintro.org. (A hard copy can be purchased on Amazon for \$25 or less.). Referred to in the syllabus as *DBR*.

There will also be additional readings available on the course webpage.

Grading

Your course grade will be calculated as follows:

Discussion Leader 10%

Problem Sets 30%

Group project 30%

Midterm exam 30%

Letter Grade	Score (% Correct)
A	90-100%
A-	88-89%
B+	84-87%
B	80-83%
B-	78-79%
C+	74-77%
C	67-73%
D	60-66
F	59 and below

Attendance is mandatory and class participation and class discussion is an essential part of this class.

Discussion Leader

During the course of the class there will be opportunities to discuss a paper related to a topic that we are working on or to present solutions to one of the 6 problem sets to your fellow students.

Participation in either of these activities will count toward your attendance/participation grade.

Discussion leaders roles: During the semester there are a number of papers assigned for reading that are related to the content that we are discussing that week. As part of your attendance and participation grade, you must sign up to be a discussion leader for **ONE** paper.

Your role as a discussion leader will be to give a 10-15 minute summary of the paper and pose some discussion questions for the class to talk about.

R Labs

Throughout the semester, we will devote about 15-30 minutes of class time each week to getting acquainted with the statistical package **R**. During these labs, I will introduce a programming topic and then ask you to complete a short assignment in breakout groups of 3 or fewer students.

In preparation for the **R** Labs, please download and install:

- 1) The latest version of **R**: <https://cloud.r-project.org/>
- 2) The latest version of **RStudio**:
<https://www.rstudio.com/products/rstudio/download/>.

Problem Sets

Your final grade will be based on 6 problem sets with assignments provided one week before the due date. These assignments will be taken from the required text. They must be turned in on time and late assignments will receive a grade of zero in the absence of a university-approved excuse.

Take-Home Midterm Exam

There will be one take-home exam for the class: an in-class midterm on **Thursday, October 17th**. The midterm will be about ½ multiple choice and ½ short answer questions. You will need a simple calculator that can do addition, multiplication and division for both the midterm and the final. The midterm will include pieces of R code that you will be asked to interpret as well.

Group Research Project

The group research project will require you to work in groups of 2-4 on one of a few pre-selected topics that I will provide to you in early March. Groups may also choose their own topic but you must consult with me beforehand if you decide to pursue your own topic.

The group research project will require you to apply the methods that you will learn in the class to a policy problem. There are two main deliverables for the project:

1. A 5-10 page report.
2. A 5-7 minute presentation on your findings.

You will be required to present your findings to the class and submit your reports on **Thursday, November 21st (THIS IS OUR LAST DAY OF CLASS)** There is more detailed information about the group project here: **TBA**.

The group projects will involve a final debate in the style of NPRs Intelligence Squared. Please take a look at some of these podcasts to get a sense of how these debates are hosted. Each of these debates **WILL HAVE A WINNING GROUP** that will be awarded a prize.

GROUP PROJECT SIGNUP SHEET: Please sign up to be in a group within the next two weeks using this Google Document: **TBA**

Group Project Proposal

On **October 24th**, you will be asked to submit a final group project proposal for me to review. The final group project proposal will be a 1-2 paragraph document with the following information:

1. The topic that you are proposing to work on.
2. The position that you plan to take.
3. Data/evidence that you will use to support your argument.

Key Dates

- **October 17th:** Take home midterm.
- **October 24th:** Final group project proposal due.
- **November 21st:** Group research projects and presentations due.

Course Calendar

I reserve the right to change the calendar to ensure that we spend enough time on each topic. If changes become necessary, they will be announced in class.

Week 1 August 15:

Course introduction and data concepts • DBR Ch. 1

Week 2 August 22:

Sources and nature of data • DBR Ch. 1

R Lab 1: Introduction to R: loading data.

- Read part 1 of this tutorial
<http://web.cs.ucla.edu/~gulzar/rstudio/basic-tutorial.html>

Problem Set 1 Released

Week 3 August 29:

Experiments, observational studies and causality • DBR Ch. 1

R Lab 3: Loading and manipulating data (continued).

- Part 2 of tutorial: <http://web.cs.ucla.edu/~gulzar/rstudio/basic-tutorial.html>.

Problem Set 1 Due

Week 4 September 5:

Probability and Distributions • DBR, Chapters 2 & 3

Problem Set 2 Released**Week 5 September 12:**

Foundations for inference • DBR, Chapters 4

R Lab 4: Data visualization.

- Part 4 of tutorial: <http://web.cs.ucla.edu/~gulzar/rstudio/basic-tutorial.html>

Problem Set 2 Due**Week 6 September 19:**

Inference for numerical data • DBR, Chapter 5

R Lab 5: Hypothesis testing in R.

Problem Set 3 Released**Week 7 September 26:**

Inference for categorical data • DBR, Chapter 6

R Lab 6: Hypothesis testing in R for categorical data.

Problem Set 3 Due**Week 8 October 3rd:**

Introduction to linear regression I • DBR, Chapter 7

R Lab 7: Introduction to linear regression.

Week 9 October 10th:

Introduction to linear regression II • DBR, Chapter 7

R Lab 8: Introduction to linear regression, continued.

Problem Set 4 Released

Week 10 October 17th:

Take home midterm exam

FINAL GROUP PROJECT PROPOSAL DUE: 10/24

Week 11 October 24th:

Introduction to multiple regression I • DBR, Chapter 8

R Lab 9: Multiple regression.

Problem Set 4 Due

Week 12 October 31st:

Introduction to multiple regression II • DBR, Chapter 8

R Lab 10: Multiple regression.

Problem Set 5 Released

Week 13 November 7th:

Introduction to logistic regression • DBR, Chapter 8

R Lab 11: Logistic regression.

Problem Set 5 Due

Week 14 November 14th:

Special topic: Propensity score matching • [Michael Anderson's \(Berkeley\) Lecture Notes](#)

R Lab 12: Propensity score matching.

Problem Set 6 Released

Week 15 November 21st: Group project research presentations.

FINAL GROUP PROJECTS AND PRESENTATIONS DUE: 4/30

Problem Set 6 Due

Statement about Students with Disabilities

Students with special needs that require accommodation should notify me and the Office for Disability Services in the first two weeks of the course so appropriate arrangements can be made. All information and documentation of special needs is confidential.

Statement about Plagiarism and Academic Dishonesty

Students are responsible for maintaining the highest standards of honesty and integrity in every phase of their academic careers. The penalties for academic dishonesty are severe and ignorance of the policy is not an acceptable defense. See also <https://ovpi.uga.edu/academic-honesty>.