

**ANDREW YOUNG SCHOOL OF POLICY STUDIES
DEPARTMENT OF PUBLIC MANAGEMENT & POLICY
PMAP 4041: POLICY DATA ANALYSIS
SPRING 2020**

Instructor: Yuriy Davydenko
Office: 14 Marietta St. NW, AYSPS 333G
E-mail: ydavydenko1@gsu.edu

Class Time: Mon & Wed 02:00 - 03:15 PM
01/13/2020 - 05/05/2020

Class Location: Classroom South 227

Schedule an appointment at: calendly.com/ydavydenko1

PREREQUISITE

Math 1070, minimum C grade.

COURSE DESCRIPTION

This course focuses on quantitative research methods applicable to the study of public policy. Students will be introduced to the use of descriptive statistics as well as to the development and testing of empirical hypotheses using basic inferential statistical methods.

This is a rigorous course, requiring 4-6 hours of work per week outside of class, *on average*. You are expected to read the lecture notes and any other materials provided by the instructor each week before coming to class. Remember, the key to learning is persistence.

LEARNING OBJECTIVES

At the end of the course, students should be able to:

- Understand the value of data analysis and formulate questions that can be answered using data
- Understand the structure and organization of data
- Understand data collection methods and their implications for the scope of inference
- Manipulate data and run basic statistical analyses using R software
- Conduct basic exploratory data analysis
- Calculate descriptive statistics
- Make basic data visualizations
- Use scatterplots and correlation coefficients to show the direction and strength of relationships between interval-level variables
- Construct and Interpret contingency tables
- Interpret regression coefficients in both bivariate and multiple regressions
- Justify claims based on confidence intervals and hypothesis testing
- Document and share data analyses with others

TEXTS, TOOLS, AND OTHER RESOURCES

1. iCollege

iCollege will be used for communication and to disseminate course content, lecture notes, PowerPoint slides, datasets, and more. Students are expected to use this resource on a regular basis for course materials and announcements.

2. Course Website

For your convenience, materials relevant to this course, including examples of analyses and reports, will be available at the course website <https://yuriygdv.github.io/pmap4041spring2020>

3. Required Textbooks

OpenIntro Statistics, 4th Edition , available at <https://leanpub.com/openintro-statistics>

Roger D. Peng (2016), Exploratory Data Analysis with R, available at <https://leanpub.com/exdata>

4. Recommended Textbooks

Advanced High School Statistics, Second Edition, available at <https://leanpub.com/ahss>

Moore, David S., Notz, William, I., & Fligner, Michael, A. (2011). The Basic Practice of Statistics (with Student CD). 6th Edition. W.H. Freeman. (or any newer edition)

5. For those of you who require more detailed treatment of the statistics topics, the following textbooks are recommended as additional resources (earlier editions are also fine):

Meier, Kenneth, Jeffrey Brudney, and John Bohte. Applied statistics for public and nonprofit administration. Cengage Learning, 2011.

Healey, Joseph F. (2013). The Essentials of Statistics: A Tool for Social Research. 3rd Edition. Wadsworth Cengage Learning.

Weiers, R. M. "Introduction to Business Statistics. 2005." Thomson Brooks/Cole, Belmont, CA, USA.

6. R Software: You are required to use R software for all homework assignments. R is a free software environment for statistical computing and graphics available at www.r-project.org . In this course, we will use RStudio - an open-source integrated development environment for R available at <https://rstudio.com/products/rstudio/> and RStudio Cloud at <https://rstudio.cloud/> .

7. R Books & Resources:

- **R for Data Science** by Garrett Grolemund and Hadley Wickham, at <https://r4ds.had.co.nz/>
- **An Introduction to R** by Longhow Lam (PDF, 2010-10-28, 212 pages), at https://cran.r-project.org/doc/contrib/Lam-IntroductionToR_LHL.pdf
- **Cookbook for R** at <http://www.cookbook-r.com/>
- **R Programming Tutorial** on YouTube - Learn the Basics of Statistical Computing <https://www.youtube.com/watch?v=V8eKsto3Ug>

8. Computers: For the purposes of this course, having a personal laptop with all the required software (R and RStudio) is recommended but not required. All the course assignments can be completed in RStudio Cloud using the library computers.

INTERACTIVE VIDEO TECHNOLOGY TRAINING

LinkedIn Learning (formerly Lynda.com) offers an online training library on a variety of topics, including software tools for data analysis. Service is free to Georgia State University's students. You can access online video tutorials with your CampusID and password using the following link:

<https://technology.gsu.edu/technology-services/it-services/training-and-learning-resources/linkedin-learning/>

For the purposes of this course, the following training can be very helpful:

Learning R by Barton Poulson (2h 51m Beginner + Intermediate Released: Aug. 29, 2019)

<https://www.linkedin.com/learning/learning-r-2/r-for-data-science?u=76216298>

R Statistics Essential Training by Barton Poulson (5h 59m, Beginner + Intermediate, Released: Sep. 26, 2013)

<https://www.linkedin.com/learning/r-statistics-essential-training/next-steps?u=76216298>

GRADING

Grades for this course will be based on:

Assignment	Weight
Problem Sets	20%
Computer Assignments	20%
Midterm Exam	20%
Data Analysis Project	20%
Final Exam	20%
Pop-Up Quizzes (extra points)	0-10%

Final grades will be assigned according to the following schedule:

A+	97 – 100
A	93 – 96
A-	90 – 92
B+	87 – 89
B	83 – 86
B-	80 – 82
C+	77 – 79
C	73 – 76
C-	65 – 72
D	60 – 64
F	< 60

Problem Sets

There will be up to 10 problem sets to complete over the semester. The problem sets will be based on the material from the previous class session. I will make them available on iCollege at the end of each

lecture. Typically, you will have one week to complete a problem set. Students are allowed to discuss problems set in groups but must submit their own work.

Computer Assignments

You will have up to ten computer assignments over the course of the semester. These will be posted after the in-class computer workshops. These assignments will give you a chance to apply statistical concepts and to practice using the R statistical software in conducting statistical analysis and developing statistical reports. Typically, you will have one week to complete a problem set.

Both problem sets and computer assignments should be relatively short, and you will be given ample time to complete each assignment. Please note that assignments turned in up to one week late will be worth 50% of the points earned. After one week, no credit will be given for late work. After one week, no credit will be given for late work.

All homework should be turned in an appropriate Dropbox on iCollege. I will not accept hard copies of assignments or emailed assignments at any time.

Exams

You will take a mid-term exam and a final exam in this course. Both exams will be closed-book and will be conducted in class. The content of the exams will not be cumulative, except for the fact that statistical concepts tend to build on each other. The format for both exams will be a combination of multiple-choice and short-answer questions. **There will be no make-up exams.** The only two acceptable reasons for missing an exam are a death in the immediate family or your hospitalization. In either of these cases, please notify me as soon as possible and no later than the start of the exam.

Data Analysis Project

You will have a small data analysis project where you will apply what you have learned in this course to answer a practical question. At the end of the semester, you will present your project to your peers.

Quizzes

Please expect a few short (up to 15-minute-long) quizzes during the semester. The quizzes will be closed-book and will be conducted at the beginning of class. The quizzes' content will be based on the material from the previous lectures. There will be no make-up quizzes.

COURSE POLICIES

Attendance Policy

I will not grade attendance in this course. However, it will be very difficult for you to succeed in this course if you do not attend lectures. As such, I expect students to attend all lectures and to arrive on time. I will take attendance at the start of every class for statistical purposes.

Statement on Academic Honesty

You are expected to abide by Georgia State University's regulations on academic honesty. These regulations may be found in the *Georgia State University Undergraduate Catalog 2014-15* (<http://catalog.gsu.edu/undergraduate20142015/university-academic-regulations/>). Examples of

academic dishonesty include (but are not limited to) plagiarism, cheating on examinations, unauthorized collaboration, falsification and multiple submission. Please take the time to familiarize yourself with the University's Academic Honesty Policy. You are reminded that, for this course, every piece of work submitted must be your own. Academic dishonesty will result in a grade of "F" on that piece of work and likely in the course as well. Please be aware that violations of the Academic Honesty Policy may also result in your dismissal from the University.

Incompletes and Withdrawals

A grade of "Incomplete" will only be given in RARE circumstances, namely in the event of a death in the immediate family or your hospitalization.

If you wish to withdraw from the course, you must do so in accordance with the University's procedures. Students who wish to avoid being given a grade of "WF" must ensure that they officially withdraw from the course before the semester mid-point. Please refer to the Office of the Registrar for further information with respect to voluntary withdrawals and deadlines.

Special Accommodations

If you have a disability and require accommodation, please speak with me on the first day of class. You will need to register with the Office of Disability Services. In order to receive accommodation for your disability, you must also submit to me a signed Accommodation Plan issued by the Office of Disability Services within the **first two weeks of class**.

Cell Phone Policy

As a courtesy to the instructor and other students, please turn off your cell phones during lectures.

Course Evaluation

Your constructive assessment of this course plays an indispensable role in shaping education at Georgia State. Upon completing the course, please take time to fill out the online course evaluation and provide your constructive feedback.

COURSE SCHEDULE AND ASSIGNMENTS

The course syllabus provides a general plan for the course; **deviations may be necessary.**

WEEK	DATE	TOPIC	HOMEWORK ASSIGNMENT	DUE DATES & NOTES
1	Mon, Jan 13	About the Course & Intro to Data	Reading: OpenIntro: Ch. 1, pp.1-16 Applying Concepts: Problem Set 1	
	Wed, Jan 15	R primer	R: Computer Assignment 1 (Playing with R: Swirl & RStudio Cloud Primers)	
2	Mon, Jan 20	MLK, No Class	---	
	Wed, Jan 22	Intro to Data, Basic Statistical Concepts	Reading: OpenIntro Statistics, Ch. 1, pp.16-34 Applying Concepts: Problem Set 2	Problem Set 1
3	Mon, Jan 27	Summarizing Data: Frequency Distributions; Graphing Frequency Distributions	UNIVARIATE ANALYSIS	
			Reading: Lecture Notes 3 Applying Concepts: Problem Set 3	Computer Assignment 1
	Wed, Jan 29		R: Computer Assignment 2	Problem Set 2
4	Mon, Feb 3	Summarizing Data: Describing Frequency Distributions with Numbers	Reading: OpenIntro, Ch.2 (pp. 42-52); Lecture Notes 4 Applying Concepts: Problem Set 4	Problem Set 3
	Wed, Feb 5	Summarizing Data	R: Computer Assignment 3	Computer Assignment 2
		BIVARIATE ANALYSIS		
5	Mon, Feb 10	Scatterplots & Correlations	Reading: OpenIntro, Ch.2.1 (pp. 41-42) & Ch. 8.1.4 (p.310-311); Lecture Notes 5 Applying Concepts: Problem Set 5	Problem Set 4
	Wed, Feb 12	Scatterplots & Correlations	R: Computer Assignment 4	Computer Assignment 3
6	Mon, Feb 17	Bivariate Regression	Reading: Lecture Notes 6; OpenIntro, Ch 8.1 (pp.305-310) & Ch. 8.2 (pp. 317-324) Applying Concepts: Problem Set 6	Problem Set 5
	Wed, Feb 19	Bivariate Regression	R: Computer Assignment 5	Computer Assignment 4
7	Mon, Feb 24	Contingency Tables	Reading: OpenIntro, Ch 2.2 (pp.61-68); Lecture Notes 7 Applying Concepts: Problem Set 7	Problem Set 6
	Wed, Feb 26	Contingency Tables	R: Computer Assignment 6	Computer Assignment 5
8	Mon, Mar 2	MIDTERM EXAM (Mar 3 - Last day to WTHDR)		Problem Set 7
	Wed, Mar 4	REVIEW		Computer Assignment 6
9	Mon, Mar 9	Multiple Regression	Reading: Lecture Notes 8 ; OpenIntro, Ch 9.1 - 9.2 (pp.343-356) Applying Concepts: Problem Set 8	
	Wed, Mar 11	Multiple Regression	R: Computer Assignment 7	
	Mon, Mar 16	Spring Break - Mar 16-22	No Class	

Exploratory Data Analysis/Descriptive Statistics

	Wed, Mar 18	Spring Break - Mar 16-22	No Class	
10	Mon, Mar 23	Probability & Normal Distributions	Reading: OpenIntro, Ch.3 Probability (recommended but not required), Ch. 4.1 Normal Distribution Applying Concepts: Problem Set 9	Problem Set 8
	Wed, Mar 25	Probability & Normal Distributions	R: Computer Assignment 8	Computer Assignment 7
11	Mon, Mar 30	Foundations of Inference: Sampling Distributions, Confidence Intervals, Hypothesis Testing	Reading: OpenIntro, Ch. 5 Applying Concepts: Problem Set 10	Problem Set 9
	Wed, Apr 1	Inference for Categorical Data	Reading: OpenIntro, Ch. 6 R: Computer Assignment 9	Computer Assignment 8
12	Mon, Apr 6	Inference for Numerical Data	Reading: OpenIntro, Ch. 7.1-7.3 Applying Concepts:	
	Wed, Apr 8	Inference for Linear Regression	Reading: OpenIntro, Ch. 8.4 R: Computer Assignment 10	Computer Assignment 9
13	Mon, Apr 13	Applying Concepts	Working on the project	
	Wed, Apr 15	Applying Concepts (Tax Day)	Working on the project	Problem Set 10
14	Mon, Apr 20	Project Presentations		Computer Assignment 10
	Wed, Apr 22	Project Presentations		
15	Mon, Apr 27 (Last Class)	Review	OpenIntro, Lecture Notes, Lecture Slides, Problem Sets	
	Apr 28 - May 5	Final Exams		
	May 7	Grades Due		

From Exploration to Inference