

PADP7120-26207: Data Applications in Public Administration

Department of Public Administration & Policy
School of Public & International Affairs
The University of Georgia

Spring 2020: Thursdays @ 3:30pm – 6:15pm
Caldwell 306

Instructor: Alex Combs, Ph.D.

Office: Baldwin Hall 278

Office Hours: Thursday 11:00am – 2:00pm

E-mail: Alex.Combs@uga.edu

Course Overview

This course introduces graduate students of public administration and policy to key concepts in data applications, emphasizing both statistical theory and software skills necessary to understand and perform quantitative analysis.

This course is organized into three broad sections:

1. Fundamental statistical concepts, data description, and visualization.
2. Inference: How sample statistics are used to explain facts & events in a population.
3. Regression: The workhorse of predictive and explanatory statistical analysis.

Learning Objectives

Student progress on learning objectives will be assessed via in-class labs, problem sets, and exams. Upon successful completion of this course, students should be able to:

1. Identify and resolve issues often encountered during quantitative analysis.
2. Identify the appropriate methodological tools for analysis.
3. Analyze and interpret data using the methodologies covered in the course.
4. Produce data analysis to effectively inform the public and other stakeholders.

Course Materials

All texts and software for this course are open source or available to you through UGA.

Texts

1. Grolemund, Garrett, and Wickham, Hadley. 2017. R for Data Science. O'Reilly. <https://r4ds.had.co.nz>. Denoted as **G&W** in the Course Schedule.
2. Ismay, Chester, and Kim, Albert Y. 2019. Statistical Inference via Data Science. <https://moderndive.com>. Denoted as **I&K** in the Course Schedule.

Software

1. R <https://cran.r-project.org/>
 - R is the program that executes computational commands.
2. RStudio <https://www.rstudio.com>
 - RStudio is a user-interface for R. **Install R before installing RStudio.**

Other potentially useful texts

1. Lower-level statistics review: https://www.openintro.org/stat/index.php?stat_book=aps.
2. Ismay, Chester. 2018. Getting Used to R, RStudio, and RMarkdown. <https://rbasics.netlify.com/index.html>
3. Diez, David M., Christopher D. Barr, and Mine ĀĖetinkaya-RĀijndel. 2016. OpenIntro Statistics. 3rd Eds. https://www.openintro.org/stat/index.php?stat_book=os.
4. Healy, Kieran. 2018. Data Visualization: A practical introduction. Princeton University Press. <http://socviz.co/index.html#preface>.

Additional readings from other sources will be distributed in class or online.

Assignments & Grading

Course grades are determined by performance on the following assignments:

- In-Class Labs – 10%
- DataCamp Chapters – 20%
- 4 Problem Sets – 40%
- Midterm Exam – 15%
- Final Exam – 15%

In-Class Labs

- About half of our meetings will include an in-class lab to practice covered concepts.
- Labs are graded as complete/incomplete. If you attend class and work on a lab, then you receive full credit.

DataCamp Chapters

- <https://www.datacamp.com>
- A variety of DataCamp chapters will be assigned to provide an opportunity to practice R between class meetings.
- DataCamp assesses your work immediately. Your course grade on this component will incorporate your performance on DataCamp.

Problem Sets

- Problem sets include a combination of conceptual and applied questions. Applied questions require students to execute statistical operations using R.
- Up to 3 students may work together on a problem set.
- Students should submit their answers to problem sets as well as their R script.

Exams

- Both midterm and final exams will be distributed online.
- You are free to use notes, texts, and any other resources.
- They will be timed.
- Both will include questions that cover concepts and require analysis in R.

Course & University Policies

Instructor Availability

If you would like to speak with me outside of class, feel free to stop by my office during scheduled office hours. If you need to meet outside of office hours, please email me at least 24 hours in advance to set a specific time.

Class Attendance/Participation

Students are highly encouraged to participate through asking questions and answering inquiries raised in class. The readings are curated to help you prepare for class, practice work relevant to assignments, and inspire questions. You are highly encouraged to complete the readings. Not all assigned readings may be discussed or used in class.

Special Needs and Academic Honesty

Students with special needs that require accommodation should notify the Office for Disability Services and me as soon as possible so the appropriate arrangements can be made. All information as well as documentation is considered confidential. UGA Student Honor Code: "I will be academically honest in all of my academic work and will not tolerate academic dishonesty of others." A Culture of Honesty, the University's policy and procedures for handling cases of suspected dishonesty, can be found at www.uga.edu/ovpi. Students are responsible for informing themselves about those standards before performing any academic work.

Miscellaneous

No extra credit will be assigned in this course under any circumstances. Keep in mind that final grades may only be changed in the event of a clerical error (e.g., points summed incorrectly). Also, for privacy reasons, information pertaining to course grades cannot be discussed over the telephone or via email. A final grade of Incomplete will only be given in this course under extraordinary circumstances and is solely at the discretion of the instructor. The course syllabus is a general plan for the course; deviations announced to the class by the instructor may be necessary. Please silence devices while in class. I accept late work after the due date only by prior arrangement.

PADP 7120 Spring 2020 Course Schedule

Week	Topic & Labs	Readings	Problem Sets	DataCamps
1/9	Intro to Stats RLab 1 – Intro to R	I&K Ch. 1 G&W Ch. 4, 6, 8 R Demo 1		
1/16	Data Types & Structures RLab 2 – Intro to Data Viz	I&K Ch. 2.1-2.3, 2.5, 2.8	PS #1 Assigned	1 – Intro to Basics 2 – Data Frames 3 – Language of Data
1/23	RLab 3 – Import, Wrangle, & Tidy	I&K Ch. 4.1 G&W Ch. 5, 12		4 – Exploring Raw Data 5 – Data Wrangling 6 – Tidying Data
1/30	Data Description RLab 4 – Describing Data			7 – Categorical Data 8 – Numerical Summary
2/6	Data Viz Best Practices RLab 5 – Data Viz for Communication	I&K Ch. 2 G&W Ch. 28	PS #1 Due PS #2 Assigned	9 – Data Visualization 10 – Types of Data Viz
2/13	Simple Linear Regression Multiple Regression	I&K Ch. 5, 6		11 - Correlation 12 – Simple Linear Reg 13 – Interpreting Reg
2/20	Regression with Categorical Variables Regression with Interactions			14 – Parallel Slopes 15 – Ext Par Slopes
2/27	Regression with Nonlinear Variables Model Fit		PS #2 Due	
3/5	Midterm Exam – No Class			
3/12	Spring Break – No Class			
3/19	Estimation and Inference RLab 6 – Joining Data in R	I&K Ch. 7, 8	PS #3 Assigned	
3/26	Hypothesis Testing Regression Inference	I&K Ch. 9, 10		
4/2	Regression Assumptions Regression Diagnostics	Kabacoff Ch. 8		
4/9	Causation & Bias RLab 7 – Omitted Variable Bias	Pearl Ch. 1	PS #3 Due PS #4 Assigned	
4/16	Intro to Panel Data Analysis			
4/23	Intro to Time Series RLab 8 - Forecasting		PS #4 Due	
4/30	Final Exam – No Class			